



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 04:17 am GMT

PDB ID : 3WEZ  
Title : Crystal structure of human beta-galactosidase in complex with NOEV  
Authors : Suzuki, H.; Ohto, U.; Shimizu, T.  
Deposited on : 2013-07-16  
Resolution : 2.11 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

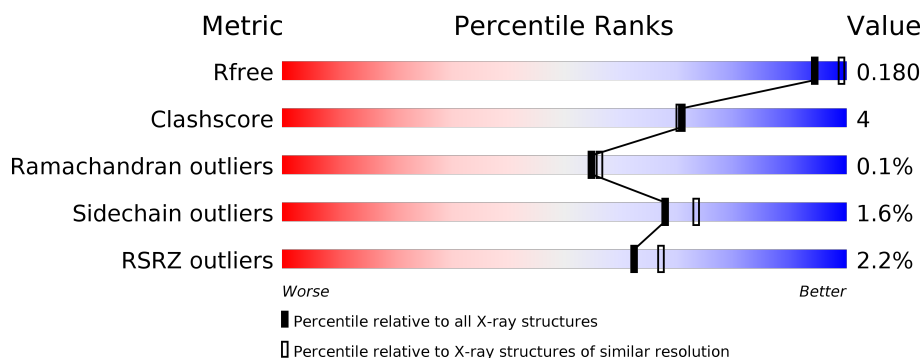
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.11 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))
$R_{free}$	100719	4988 (2.14-2.10)
Clashscore	112137	5557 (2.14-2.10)
Ramachandran outliers	110173	5504 (2.14-2.10)
Sidechain outliers	110143	5505 (2.14-2.10)
RSRZ outliers	101464	5021 (2.14-2.10)

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. 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	678	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 3%, orange 11%, yellow 80%, green 89%, grey 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>3%</span> <span>80%</span> <span>9%</span> <span>11%</span> </div> </div>
1	B	678	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 3%, orange 11%, yellow 81%, green 91%, grey 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>3%</span> <span>81%</span> <span>8%</span> <span>11%</span> </div> </div>
1	C	678	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 3%, orange 11%, yellow 81%, green 91%, grey 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>3%</span> <span>81%</span> <span>8%</span> <span>11%</span> </div> </div>
1	D	678	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 3%, orange 11%, yellow 79%, green 90%, grey 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>3%</span> <span>79%</span> <span>10%</span> <span>11%</span> </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
2	NAG	A	703	-	-	-	X
2	NAG	B	703	-	-	-	X
4	N8V	A	706	-	-	-	X
4	N8V	B	706	-	-	-	X
4	N8V	C	706	-	-	-	X
4	N8V	D	706	-	-	-	X
5	SO4	A	708	-	-	-	X
5	SO4	B	708	-	-	-	X
6	EDO	A	709	-	-	-	X
6	EDO	B	710	-	-	-	X
6	EDO	C	709	-	-	X	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 21184 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	605	Total	C	N	O	S	0	9	0
			4842	3139	798	887	18			
1	B	605	Total	C	N	O	S	0	6	0
			4829	3133	795	884	17			
1	C	603	Total	C	N	O	S	0	6	0
			4810	3122	790	881	17			
1	D	602	Total	C	N	O	S	0	10	0
			4810	3118	791	883	18			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLU	-	EXPRESSION TAG	UNP P16278
A	1	ALA	-	EXPRESSION TAG	UNP P16278
A	2	GLU	-	EXPRESSION TAG	UNP P16278
A	3	ALA	-	EXPRESSION TAG	UNP P16278
A	4	TYR	-	EXPRESSION TAG	UNP P16278
A	5	VAL	-	EXPRESSION TAG	UNP P16278
A	6	GLU	-	EXPRESSION TAG	UNP P16278
A	7	PHE	-	EXPRESSION TAG	UNP P16278
A	8	HIS	-	EXPRESSION TAG	UNP P16278
A	9	HIS	-	EXPRESSION TAG	UNP P16278
A	10	HIS	-	EXPRESSION TAG	UNP P16278
A	11	HIS	-	EXPRESSION TAG	UNP P16278
A	12	HIS	-	EXPRESSION TAG	UNP P16278
A	13	HIS	-	EXPRESSION TAG	UNP P16278
A	14	ASP	-	EXPRESSION TAG	UNP P16278
A	15	TYR	-	EXPRESSION TAG	UNP P16278
A	16	LYS	-	EXPRESSION TAG	UNP P16278
A	17	ASP	-	EXPRESSION TAG	UNP P16278
A	18	ASP	-	EXPRESSION TAG	UNP P16278
A	19	ASP	-	EXPRESSION TAG	UNP P16278
A	20	ASP	-	EXPRESSION TAG	UNP P16278

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	21	LYS	-	EXPRESSION TAG	UNP P16278
A	22	THR	-	EXPRESSION TAG	UNP P16278
A	23	SER	-	EXPRESSION TAG	UNP P16278
B	0	GLU	-	EXPRESSION TAG	UNP P16278
B	1	ALA	-	EXPRESSION TAG	UNP P16278
B	2	GLU	-	EXPRESSION TAG	UNP P16278
B	3	ALA	-	EXPRESSION TAG	UNP P16278
B	4	TYR	-	EXPRESSION TAG	UNP P16278
B	5	VAL	-	EXPRESSION TAG	UNP P16278
B	6	GLU	-	EXPRESSION TAG	UNP P16278
B	7	PHE	-	EXPRESSION TAG	UNP P16278
B	8	HIS	-	EXPRESSION TAG	UNP P16278
B	9	HIS	-	EXPRESSION TAG	UNP P16278
B	10	HIS	-	EXPRESSION TAG	UNP P16278
B	11	HIS	-	EXPRESSION TAG	UNP P16278
B	12	HIS	-	EXPRESSION TAG	UNP P16278
B	13	HIS	-	EXPRESSION TAG	UNP P16278
B	14	ASP	-	EXPRESSION TAG	UNP P16278
B	15	TYR	-	EXPRESSION TAG	UNP P16278
B	16	LYS	-	EXPRESSION TAG	UNP P16278
B	17	ASP	-	EXPRESSION TAG	UNP P16278
B	18	ASP	-	EXPRESSION TAG	UNP P16278
B	19	ASP	-	EXPRESSION TAG	UNP P16278
B	20	ASP	-	EXPRESSION TAG	UNP P16278
B	21	LYS	-	EXPRESSION TAG	UNP P16278
B	22	THR	-	EXPRESSION TAG	UNP P16278
B	23	SER	-	EXPRESSION TAG	UNP P16278
C	0	GLU	-	EXPRESSION TAG	UNP P16278
C	1	ALA	-	EXPRESSION TAG	UNP P16278
C	2	GLU	-	EXPRESSION TAG	UNP P16278
C	3	ALA	-	EXPRESSION TAG	UNP P16278
C	4	TYR	-	EXPRESSION TAG	UNP P16278
C	5	VAL	-	EXPRESSION TAG	UNP P16278
C	6	GLU	-	EXPRESSION TAG	UNP P16278
C	7	PHE	-	EXPRESSION TAG	UNP P16278
C	8	HIS	-	EXPRESSION TAG	UNP P16278
C	9	HIS	-	EXPRESSION TAG	UNP P16278
C	10	HIS	-	EXPRESSION TAG	UNP P16278
C	11	HIS	-	EXPRESSION TAG	UNP P16278
C	12	HIS	-	EXPRESSION TAG	UNP P16278
C	13	HIS	-	EXPRESSION TAG	UNP P16278
C	14	ASP	-	EXPRESSION TAG	UNP P16278

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	15	TYR	-	EXPRESSION TAG	UNP P16278
C	16	LYS	-	EXPRESSION TAG	UNP P16278
C	17	ASP	-	EXPRESSION TAG	UNP P16278
C	18	ASP	-	EXPRESSION TAG	UNP P16278
C	19	ASP	-	EXPRESSION TAG	UNP P16278
C	20	ASP	-	EXPRESSION TAG	UNP P16278
C	21	LYS	-	EXPRESSION TAG	UNP P16278
C	22	THR	-	EXPRESSION TAG	UNP P16278
C	23	SER	-	EXPRESSION TAG	UNP P16278
D	0	GLU	-	EXPRESSION TAG	UNP P16278
D	1	ALA	-	EXPRESSION TAG	UNP P16278
D	2	GLU	-	EXPRESSION TAG	UNP P16278
D	3	ALA	-	EXPRESSION TAG	UNP P16278
D	4	TYR	-	EXPRESSION TAG	UNP P16278
D	5	VAL	-	EXPRESSION TAG	UNP P16278
D	6	GLU	-	EXPRESSION TAG	UNP P16278
D	7	PHE	-	EXPRESSION TAG	UNP P16278
D	8	HIS	-	EXPRESSION TAG	UNP P16278
D	9	HIS	-	EXPRESSION TAG	UNP P16278
D	10	HIS	-	EXPRESSION TAG	UNP P16278
D	11	HIS	-	EXPRESSION TAG	UNP P16278
D	12	HIS	-	EXPRESSION TAG	UNP P16278
D	13	HIS	-	EXPRESSION TAG	UNP P16278
D	14	ASP	-	EXPRESSION TAG	UNP P16278
D	15	TYR	-	EXPRESSION TAG	UNP P16278
D	16	LYS	-	EXPRESSION TAG	UNP P16278
D	17	ASP	-	EXPRESSION TAG	UNP P16278
D	18	ASP	-	EXPRESSION TAG	UNP P16278
D	19	ASP	-	EXPRESSION TAG	UNP P16278
D	20	ASP	-	EXPRESSION TAG	UNP P16278
D	21	LYS	-	EXPRESSION TAG	UNP P16278
D	22	THR	-	EXPRESSION TAG	UNP P16278
D	23	SER	-	EXPRESSION TAG	UNP P16278

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	D	1	Total	C	N	O	0	0
			14	8	1	5		
2	D	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

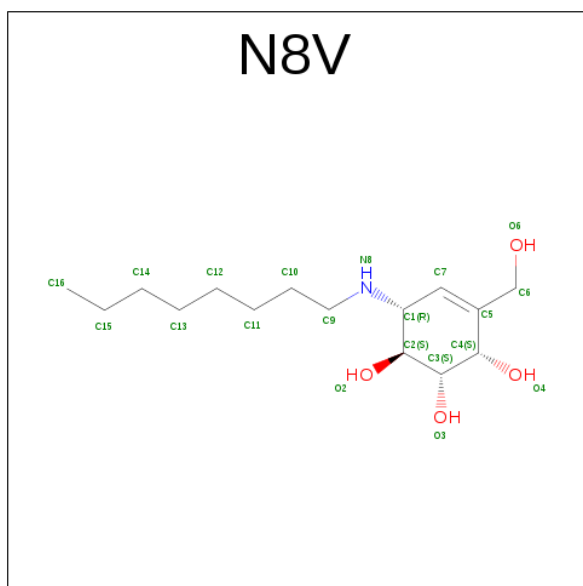
*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	D	1	Total	C	N	O	0	0
			14	8	1	5		
2	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cl	0	0
			1	1		
3	A	1	Total	Cl	0	0
			1	1		
3	D	1	Total	Cl	0	0
			1	1		
3	C	1	Total	Cl	0	0
			1	1		

- Molecule 4 is (1S,2S,3S,6R)-4-(HYDROXYMETHYL)-6-(OCTYLAMINO)CYCLOHEX-4-ENE-1,2,3-TRIOL (three-letter code: N8V) (formula: C<sub>15</sub>H<sub>29</sub>NO<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			20	15	1	4		
4	B	1	Total	C	N	O	0	0
			20	15	1	4		
4	C	1	Total	C	N	O	0	0
			20	15	1	4		

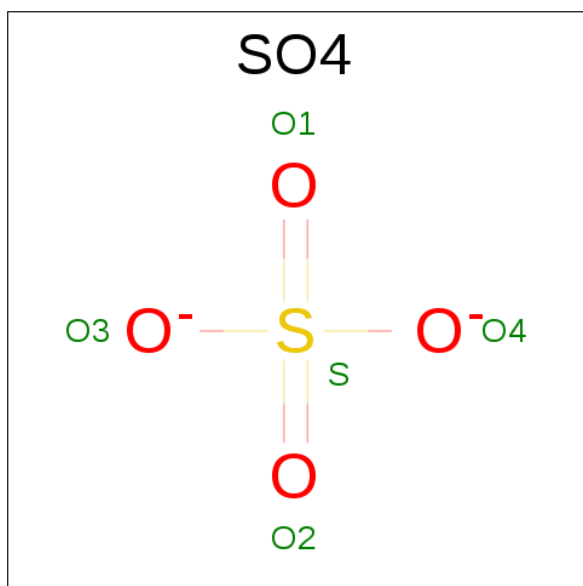
*Continued on next page...*



*Continued from previous page...*

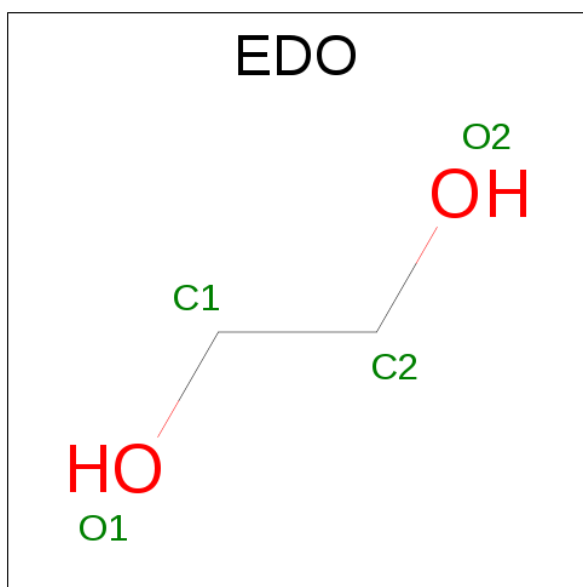
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	D	1	Total	C	N	O	0	0
			20	15	1	4		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



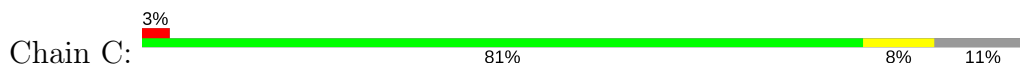
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		
6	C	1	Total	C	O	0	0
			4	2	2		
6	C	1	Total	C	O	0	0
			4	2	2		
6	D	1	Total	C	O	0	0
			4	2	2		
6	D	1	Total	C	O	0	0
			4	2	2		

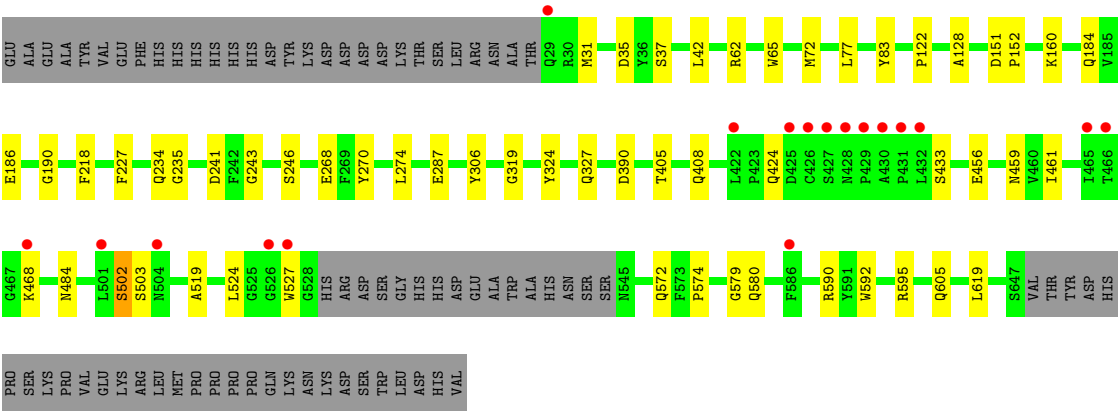
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	384	Total	O	0	0
			384	384		
7	B	412	Total	O	0	0
			412	412		
7	C	365	Total	O	0	0
			365	365		
7	D	352	Total	O	0	0
			352	352		

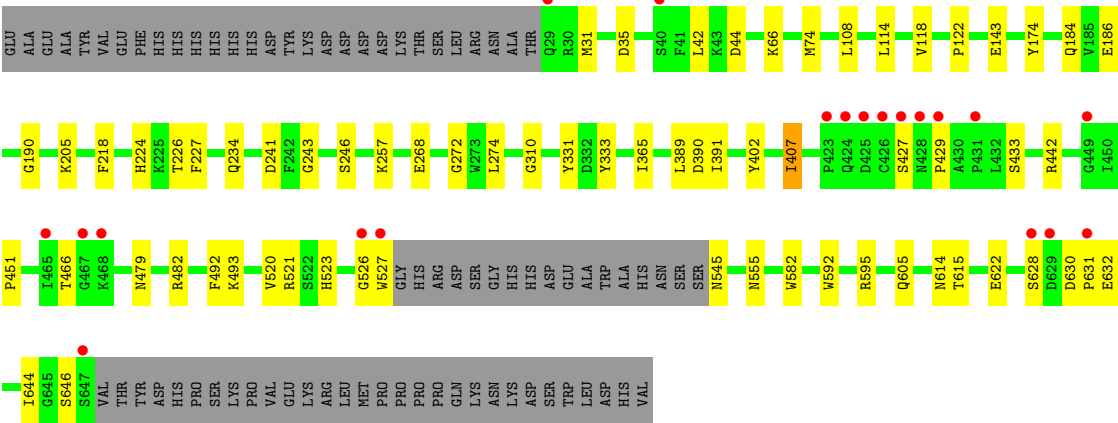
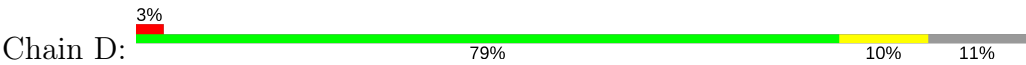


- Molecule 1: Beta-galactosidase





● Molecule 1: Beta-galactosidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.93Å 115.92Å 140.51Å 90.00° 92.25° 90.00°	Depositor
Resolution (Å)	47.43 – 2.11 47.43 – 2.11	Depositor EDS
% Data completeness (in resolution range)	98.2 (47.43-2.11) 98.2 (47.43-2.11)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.69 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.179 , 0.225 0.181 , 0.180	Depositor DCC
$R_{free}$ test set	8600 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	20.4	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 48.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.024 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	21184	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.71% 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: SO4, N8V, EDO, NAG, CL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.65	0/5039	0.72	0/6873
1	B	0.66	0/5010	0.73	2/6835 (0.0%)
1	C	0.63	0/4990	0.72	0/6809
1	D	0.62	0/5014	0.72	0/6840
All	All	0.64	0/20053	0.72	2/27357 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	640	ASP	CB-CG-OD1	5.19	122.97	118.30
1	B	568	ASP	CB-CG-OD2	-5.12	113.69	118.30

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	A	4842	0	4702	43	0
1	B	4829	0	4693	37	0
1	C	4810	0	4678	36	0
1	D	4810	0	4675	47	0
2	A	56	0	52	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	56	0	52	0	0
2	C	56	0	52	0	0
2	D	56	0	52	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	20	0	29	0	0
4	B	20	0	29	0	0
4	C	20	0	29	0	0
4	D	20	0	29	0	0
5	A	10	0	0	0	0
5	B	10	0	0	0	0
5	C	10	0	0	1	0
5	D	10	0	0	0	0
6	A	8	0	12	0	0
6	B	8	0	12	3	0
6	C	8	0	12	4	0
6	D	8	0	12	2	0
7	A	384	0	0	4	0
7	B	412	0	0	6	0
7	C	365	0	0	4	0
7	D	352	0	0	5	0
All	All	21184	0	19120	158	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 4.

The worst 5 of 158 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:A:407:ILE:CD1	1:A:492:PHE:O	2.19	0.90
1:C:405:THR:H	1:C:408:GLN:HE21	1.19	0.88
1:A:437:ASN:HD21	1:A:458:ASN:H	1.12	0.88
1:A:407:ILE:HD11	1:A:492:PHE:O	1.72	0.87
1:C:243:GLY:H	6:C:709:EDO:H11	1.42	0.83

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	610/678 (90%)	585 (96%)	23 (4%)	2 (0%)	44	43
1	B	607/678 (90%)	585 (96%)	22 (4%)	0	100	100
1	C	605/678 (89%)	581 (96%)	24 (4%)	0	100	100
1	D	608/678 (90%)	581 (96%)	26 (4%)	1 (0%)	51	51
All	All	2430/2712 (90%)	2332 (96%)	95 (4%)	3 (0%)	55	56

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	628	SER
1	A	628	SER
1	A	611	SER

### 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	527/585 (90%)	518 (98%)	9 (2%)	66	71
1	B	524/585 (90%)	516 (98%)	8 (2%)	70	75
1	C	522/585 (89%)	514 (98%)	8 (2%)	70	75
1	D	526/585 (90%)	516 (98%)	10 (2%)	62	67
All	All	2099/2340 (90%)	2064 (98%)	35 (2%)	68	71

5 of 35 residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	B	605	GLN
1	C	186	GLU
1	D	605	GLN
1	B	627	SER
1	C	31	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	321	ASN
1	B	529	HIS
1	D	97	GLN
1	B	355	GLN
1	B	567	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 40 ligands modelled in this entry, 4 are monoatomic - leaving 36 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
2	NAG	A	701	1	14,14,15	1.50	2 (14%)	15,19,21	1.79	4 (26%)
2	NAG	A	702	1	14,14,15	0.49	0	15,19,21	0.86	0
2	NAG	A	703	1	14,14,15	0.83	0	15,19,21	2.85	7 (46%)
2	NAG	A	704	1	14,14,15	0.70	0	15,19,21	1.74	4 (26%)
4	N8V	A	706	-	19,20,20	0.98	2 (10%)	16,25,25	1.09	2 (12%)
5	SO4	A	707	-	4,4,4	0.39	0	6,6,6	0.32	0
5	SO4	A	708	-	4,4,4	0.51	0	6,6,6	0.53	0
6	EDO	A	709	-	3,3,3	0.31	0	2,2,2	0.96	0
6	EDO	A	710	-	3,3,3	0.66	0	2,2,2	0.42	0
2	NAG	B	701	1	14,14,15	1.35	1 (7%)	15,19,21	1.41	2 (13%)
2	NAG	B	702	1	14,14,15	1.09	1 (7%)	15,19,21	1.75	2 (13%)
2	NAG	B	703	1	14,14,15	0.44	0	15,19,21	3.22	9 (60%)
2	NAG	B	704	1	14,14,15	0.80	0	15,19,21	1.36	3 (20%)
4	N8V	B	706	-	19,20,20	1.04	1 (5%)	16,25,25	1.32	1 (6%)
5	SO4	B	707	-	4,4,4	0.41	0	6,6,6	0.11	0
5	SO4	B	708	-	4,4,4	0.50	0	6,6,6	0.74	0
6	EDO	B	709	-	3,3,3	0.39	0	2,2,2	0.94	0
6	EDO	B	710	-	3,3,3	0.45	0	2,2,2	0.63	0
2	NAG	C	701	1	14,14,15	0.56	0	15,19,21	1.71	3 (20%)
2	NAG	C	702	1	14,14,15	0.86	0	15,19,21	1.88	3 (20%)
2	NAG	C	703	1	14,14,15	1.05	1 (7%)	15,19,21	1.46	2 (13%)
2	NAG	C	704	1	14,14,15	1.08	0	15,19,21	1.38	2 (13%)
4	N8V	C	706	-	19,20,20	0.65	0	16,25,25	1.04	1 (6%)
5	SO4	C	707	-	4,4,4	0.45	0	6,6,6	0.20	0
5	SO4	C	708	-	4,4,4	0.47	0	6,6,6	0.67	0
6	EDO	C	709	-	3,3,3	0.44	0	2,2,2	0.68	0
6	EDO	C	710	-	3,3,3	0.60	0	2,2,2	0.27	0
2	NAG	D	701	1	14,14,15	0.27	0	15,19,21	0.57	0
2	NAG	D	702	1	14,14,15	0.86	0	15,19,21	1.37	2 (13%)
2	NAG	D	703	1	14,14,15	0.46	0	15,19,21	2.12	2 (13%)
2	NAG	D	704	1	14,14,15	1.16	0	15,19,21	1.51	2 (13%)
4	N8V	D	706	-	19,20,20	0.78	0	16,25,25	1.05	1 (6%)
5	SO4	D	707	-	4,4,4	0.43	0	6,6,6	0.35	0
5	SO4	D	708	-	4,4,4	0.41	0	6,6,6	0.52	0
6	EDO	D	709	-	3,3,3	0.28	0	2,2,2	0.99	0
6	EDO	D	710	-	3,3,3	0.57	0	2,2,2	0.46	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	NAG	A	701	1	-	0/6/23/26	0/1/1/1
2	NAG	A	702	1	-	0/6/23/26	0/1/1/1
2	NAG	A	703	1	-	0/6/23/26	0/1/1/1
2	NAG	A	704	1	-	0/6/23/26	0/1/1/1
4	N8V	A	706	-	-	0/11/31/31	0/1/1/1
5	SO4	A	707	-	-	0/0/0/0	0/0/0/0
5	SO4	A	708	-	-	0/0/0/0	0/0/0/0
6	EDO	A	709	-	-	0/1/1/1	0/0/0/0
6	EDO	A	710	-	-	0/1/1/1	0/0/0/0
2	NAG	B	701	1	-	0/6/23/26	0/1/1/1
2	NAG	B	702	1	-	0/6/23/26	0/1/1/1
2	NAG	B	703	1	-	0/6/23/26	0/1/1/1
2	NAG	B	704	1	-	0/6/23/26	0/1/1/1
4	N8V	B	706	-	-	0/11/31/31	0/1/1/1
5	SO4	B	707	-	-	0/0/0/0	0/0/0/0
5	SO4	B	708	-	-	0/0/0/0	0/0/0/0
6	EDO	B	709	-	-	0/1/1/1	0/0/0/0
6	EDO	B	710	-	-	0/1/1/1	0/0/0/0
2	NAG	C	701	1	-	0/6/23/26	0/1/1/1
2	NAG	C	702	1	-	0/6/23/26	0/1/1/1
2	NAG	C	703	1	-	0/6/23/26	0/1/1/1
2	NAG	C	704	1	-	0/6/23/26	0/1/1/1
4	N8V	C	706	-	-	0/11/31/31	0/1/1/1
5	SO4	C	707	-	-	0/0/0/0	0/0/0/0
5	SO4	C	708	-	-	0/0/0/0	0/0/0/0
6	EDO	C	709	-	-	0/1/1/1	0/0/0/0
6	EDO	C	710	-	-	0/1/1/1	0/0/0/0
2	NAG	D	701	1	-	0/6/23/26	0/1/1/1
2	NAG	D	702	1	-	0/6/23/26	0/1/1/1
2	NAG	D	703	1	-	0/6/23/26	0/1/1/1
2	NAG	D	704	1	-	0/6/23/26	0/1/1/1
4	N8V	D	706	-	-	0/11/31/31	0/1/1/1
5	SO4	D	707	-	-	0/0/0/0	0/0/0/0
5	SO4	D	708	-	-	0/0/0/0	0/0/0/0
6	EDO	D	709	-	-	0/1/1/1	0/0/0/0
6	EDO	D	710	-	-	0/1/1/1	0/0/0/0

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	701	NAG	O5-C1	-3.12	1.38	1.43

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	701	NAG	C2-N2	-2.96	1.41	1.46
2	B	701	NAG	C2-N2	-2.76	1.41	1.46
2	C	703	NAG	O7-C7	-2.34	1.17	1.23
2	B	702	NAG	C2-N2	-2.31	1.42	1.46

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	703	NAG	O5-C1-C2	-5.29	104.11	111.47
2	B	703	NAG	C4-C3-C2	-4.82	103.95	111.02
2	B	703	NAG	O5-C1-C2	-4.38	105.38	111.47
2	B	703	NAG	O3-C3-C4	-4.29	101.02	110.36
2	C	703	NAG	O5-C1-C2	-3.97	105.95	111.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	709	EDO	3	0
5	C	708	SO4	1	0
6	C	709	EDO	4	0
2	D	702	NAG	1	0
2	D	703	NAG	1	0
6	D	709	EDO	2	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	605/678 (89%)	-0.28	7 (1%) 79 82	11, 19, 35, 92	0
1	B	605/678 (89%)	-0.30	7 (1%) 79 82	10, 18, 33, 92	0
1	C	603/678 (88%)	-0.07	18 (2%) 51 57	11, 19, 44, 96	0
1	D	602/678 (88%)	-0.08	20 (3%) 47 54	10, 20, 44, 79	0
All	All	2415/2712 (89%)	-0.18	52 (2%) 62 67	10, 19, 39, 96	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	647	SER	9.0
1	C	527	TRP	8.6
1	B	646	SER	6.7
1	A	647	SER	5.4
1	A	646	SER	5.3

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	SO4	A	708	5/5	0.91	0.14	9.24	50,50,55,60	0
5	SO4	B	708	5/5	0.89	0.17	5.60	49,51,55,55	0
2	NAG	A	703	14/15	0.78	0.19	4.59	41,45,48,48	0
6	EDO	C	709	4/4	0.84	0.17	4.34	31,31,32,37	0
6	EDO	A	709	4/4	0.88	0.15	3.91	31,33,35,36	0
4	N8V	A	706	20/20	0.95	0.14	3.74	17,18,45,47	0
2	NAG	B	703	14/15	0.63	0.22	3.50	43,50,53,56	0
4	N8V	C	706	20/20	0.94	0.16	3.38	16,20,34,34	0
6	EDO	B	710	4/4	0.95	0.15	2.80	20,23,24,26	0
4	N8V	D	706	20/20	0.95	0.15	2.67	17,21,39,40	0
4	N8V	B	706	20/20	0.95	0.14	2.38	14,16,39,40	0
2	NAG	A	701	14/15	0.93	0.12	1.29	19,22,25,30	0
6	EDO	B	709	4/4	0.89	0.15	1.29	32,32,35,41	0
6	EDO	A	710	4/4	0.95	0.12	1.09	21,22,23,25	0
2	NAG	C	702	14/15	0.79	0.18	1.05	43,53,58,59	0
6	EDO	D	709	4/4	0.95	0.10	0.62	29,30,30,36	0
2	NAG	A	702	14/15	0.95	0.10	0.53	24,26,29,30	0
5	SO4	C	708	5/5	0.92	0.12	0.44	45,46,49,52	0
2	NAG	C	703	14/15	0.85	0.22	0.43	38,45,47,50	0
2	NAG	B	701	14/15	0.92	0.12	0.42	21,24,29,30	0
3	CL	B	705	1/1	1.00	0.09	-0.33	15,15,15,15	0
2	NAG	D	702	14/15	0.88	0.13	-0.34	34,39,43,43	0
2	NAG	C	704	14/15	0.96	0.09	-0.55	20,23,27,28	0
5	SO4	D	708	5/5	0.95	0.10	-0.73	48,50,51,54	0
2	NAG	D	704	14/15	0.97	0.09	-0.81	26,27,30,31	0
3	CL	C	705	1/1	1.00	0.10	-0.95	16,16,16,16	0
2	NAG	B	702	14/15	0.95	0.08	-1.07	23,27,28,33	0
3	CL	D	705	1/1	1.00	0.07	-3.22	16,16,16,16	0
3	CL	A	705	1/1	1.00	0.05	-4.99	15,15,15,15	0
2	NAG	A	704	14/15	0.75	0.20	-	51,62,76,80	0
6	EDO	C	710	4/4	0.94	0.14	-	22,23,24,26	0
2	NAG	D	701	14/15	0.77	0.39	-	78,83,85,86	0
5	SO4	B	707	5/5	0.91	0.27	-	80,81,88,90	0
2	NAG	C	701	14/15	0.77	0.47	-	72,79,81,83	0
5	SO4	D	707	5/5	0.82	0.17	-	72,72,76,78	0
6	EDO	D	710	4/4	0.93	0.15	-	24,27,29,31	0
2	NAG	D	703	14/15	0.39	0.41	-	73,86,91,93	0
5	SO4	A	707	5/5	0.89	0.28	-	64,64,67,68	0
2	NAG	B	704	14/15	0.86	0.20	-	54,57,59,60	0
5	SO4	C	707	5/5	0.94	0.24	-	72,75,77,77	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.