



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 14, 2026 – 06:13 PM JST

PDB ID : 25PT / pdb\_000025pt  
Title : Crystal structure of TsaBGL  
Authors : Nam, K.H.  
Deposited on : 2026-04-14  
Resolution : 1.65 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

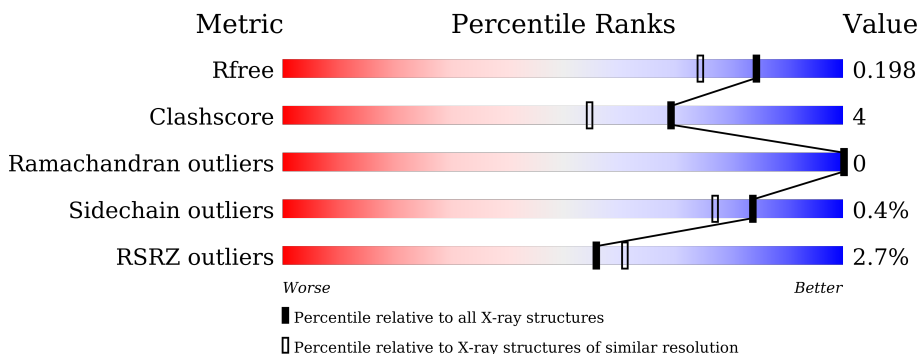
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.65 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))
$R_{free}$	180053	2563 (1.66-1.66)
Clashscore	190562	2662 (1.66-1.66)
Ramachandran outliers	187476	2621 (1.66-1.66)
Sidechain outliers	187428	2621 (1.66-1.66)
RSRZ outliers	180081	2564 (1.66-1.66)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	444	<div> <div>2%</div> <div> <div></div> <div>92%</div> <div>8%</div> </div> </div>
1	B	444	<div> <div>%</div> <div> <div></div> <div>92%</div> <div>8%</div> </div> </div>
1	C	444	<div> <div>4%</div> <div> <div></div> <div>91%</div> <div>9%</div> </div> </div>
1	D	444	<div> <div>4%</div> <div> <div></div> <div>94%</div> <div>6%</div> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 16406 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

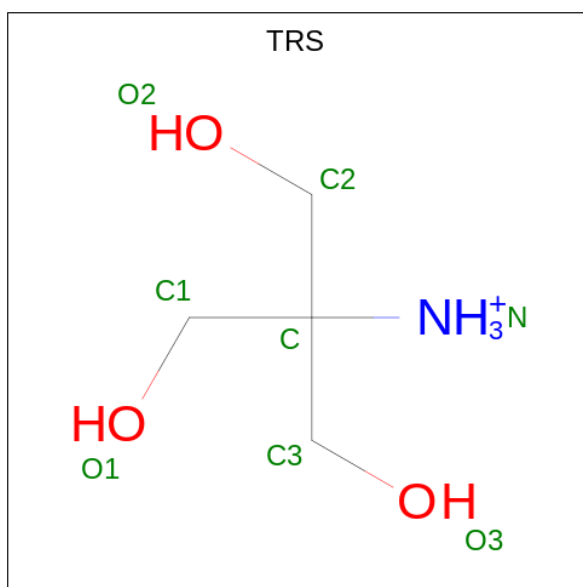
- Molecule 1 is a protein called Beta-glucosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	444	Total	C	N	O	S	0	1	0
			3673	2371	603	687	12			
1	B	443	Total	C	N	O	S	0	4	0
			3693	2384	608	689	12			
1	C	441	Total	C	N	O	S	0	0	0
			3641	2354	597	679	11			
1	D	443	Total	C	N	O	S	0	2	0
			3672	2372	602	687	11			

There are 4 discrepancies between the modelled and reference sequences:

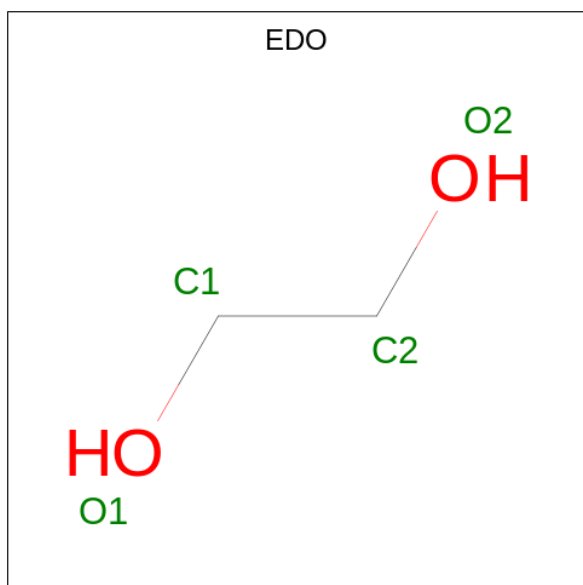
Chain	Residue	Modelled	Actual	Comment	Reference
A	2	LYS	LEU	conflict	UNP I3VXG7
B	2	LYS	LEU	conflict	UNP I3VXG7
C	2	LYS	LEU	conflict	UNP I3VXG7
D	2	LYS	LEU	conflict	UNP I3VXG7

- Molecule 2 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (CCD ID: TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			8	4	1	3		
2	B	1	Total	C	N	O	0	0
			8	4	1	3		
2	C	1	Total	C	N	O	0	0
			8	4	1	3		
2	D	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 3 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

- Molecule 4 is SODIUM ION (CCD ID: NA) (formula: Na).

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

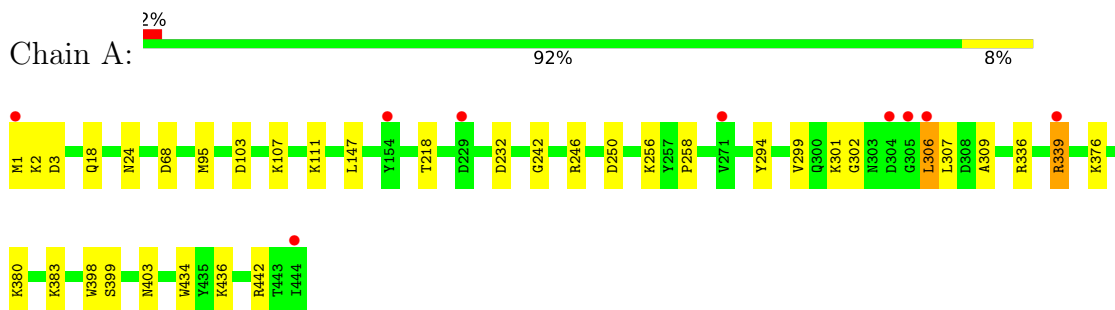
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	425	Total O 425 425	0	0
5	B	445	Total O 445 445	0	0
5	C	413	Total O 413 413	0	0
5	D	380	Total O 380 380	0	0

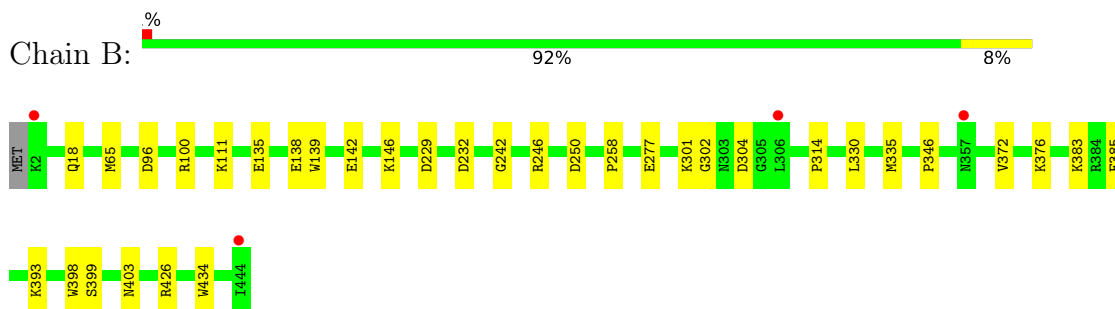
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

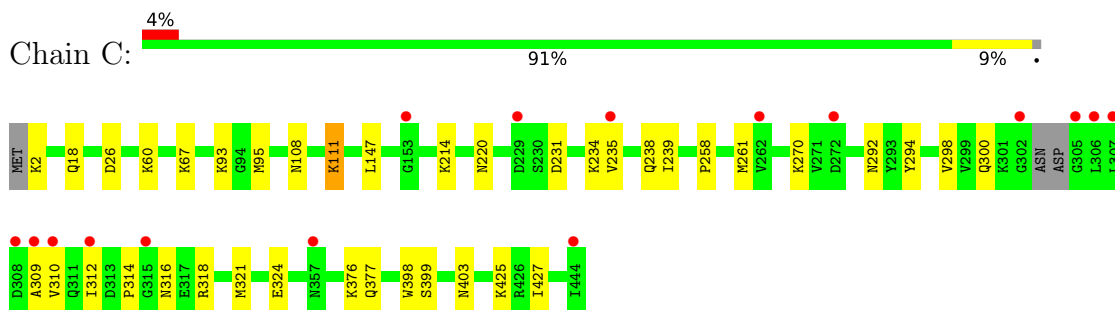
- Molecule 1: Beta-glucosidase



- Molecule 1: Beta-glucosidase

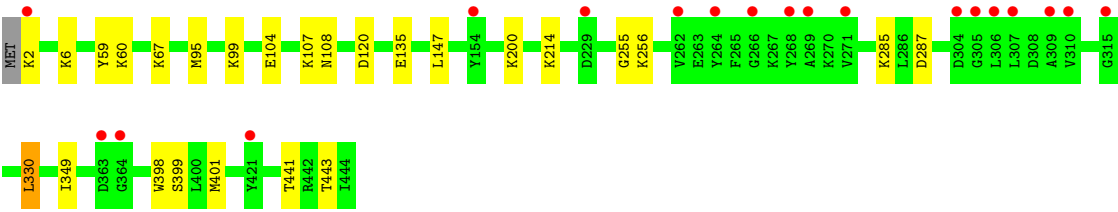


- Molecule 1: Beta-glucosidase



- Molecule 1: Beta-glucosidase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.32Å 72.77Å 97.46Å 92.49° 91.33° 95.17°	Depositor
Resolution (Å)	72.40 – 1.65 72.40 – 1.65	Depositor EDS
% Data completeness (in resolution range)	85.4 (72.40-1.65) 85.1 (72.40-1.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.26 (at 1.65Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, $R_{free}$	0.195 , 0.226 0.196 , 0.198	Depositor DCC
$R_{free}$ test set	1953 reflections (0.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.5	Xtriage
Anisotropy	0.513	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 35.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.007 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	16406	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	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 66.40 % 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 6.0787e-06. 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, TRS, EDO

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	0.16	0/3776	0.40	0/5105
1	B	0.16	0/3796	0.39	0/5131
1	C	0.15	0/3743	0.38	0/5059
1	D	0.16	0/3775	0.41	0/5105
All	All	0.16	0/15090	0.40	0/20400

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	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	339	ARG	Sidechain

### 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	3673	0	3509	24	0
1	B	3693	0	3529	24	0
1	C	3641	0	3481	28	0
1	D	3672	0	3508	25	0
2	A	8	0	12	1	0
2	B	8	0	12	0	0
2	C	8	0	12	1	0
2	D	8	0	12	0	0
3	A	4	0	6	0	0
3	B	4	0	6	0	0
3	C	8	0	12	0	0
3	D	4	0	6	0	0
4	A	4	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	425	0	0	5	1
5	B	445	0	0	4	4
5	C	413	0	0	7	3
5	D	380	0	0	5	2
All	All	16406	0	14105	101	5

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 101 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:300:GLN:HG2	1:C:312:ILE:HD11	1.63	0.79
1:D:285:LYS:HE2	1:D:285:LYS:H	1.47	0.79
1:A:256:LYS:NZ	5:A:602:HOH:O	2.18	0.73
1:D:2:LYS:N	5:D:603:HOH:O	2.26	0.69
1:D:95:MET:HG3	1:D:99:LYS:HE2	1.79	0.65

All (5) 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)
5:A:834:HOH:O	5:D:831:HOH:O[1_565]	2.14	0.06
5:B:648:HOH:O	5:C:829:HOH:O[1_665]	2.16	0.04
5:B:932:HOH:O	5:C:921:HOH:O[1_665]	2.16	0.04

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:B:788:HOH:O	5:D:732:HOH:O[1_566]	2.17	0.03
5:B:904:HOH:O	5:C:900:HOH:O[1_665]	2.19	0.01

## 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	443/444 (100%)	431 (97%)	12 (3%)	0	100	100
1	B	445/444 (100%)	434 (98%)	11 (2%)	0	100	100
1	C	437/444 (98%)	427 (98%)	10 (2%)	0	100	100
1	D	443/444 (100%)	424 (96%)	19 (4%)	0	100	100
All	All	1768/1776 (100%)	1716 (97%)	52 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	386/385 (100%)	384 (100%)	2 (0%)	81	72
1	B	388/385 (101%)	385 (99%)	3 (1%)	73	60
1	C	382/385 (99%)	381 (100%)	1 (0%)	86	80
1	D	386/385 (100%)	384 (100%)	2 (0%)	81	72
All	All	1542/1540 (100%)	1534 (100%)	8 (0%)	84	72

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	330[B]	LEU
1	D	330[A]	LEU
1	B	330[B]	LEU
1	B	330[A]	LEU
1	C	111	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	89	GLN
1	D	316	ASN
1	D	387	ASN
1	B	238	GLN
1	B	244	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 21 ligands modelled in this entry, 12 are monoatomic - leaving 9 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	TRS	B	501	-	7,7,7	0.31	0	9,9,9	0.38	0
2	TRS	D	502	-	7,7,7	0.28	0	9,9,9	0.38	0
3	EDO	C	502	-	3,3,3	0.45	0	2,2,2	0.42	0
3	EDO	D	503	-	3,3,3	0.44	0	2,2,2	0.43	0
3	EDO	A	502	-	3,3,3	0.46	0	2,2,2	0.25	0
3	EDO	B	502	-	3,3,3	0.47	0	2,2,2	0.37	0
3	EDO	C	503	-	3,3,3	0.48	0	2,2,2	0.30	0
2	TRS	C	501	-	7,7,7	0.29	0	9,9,9	0.39	0
2	TRS	A	501	-	7,7,7	0.32	0	9,9,9	0.49	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	TRS	B	501	-	-	2/9/9/9	-
2	TRS	D	502	-	-	0/9/9/9	-
3	EDO	C	502	-	-	1/1/1/1	-
3	EDO	D	503	-	-	0/1/1/1	-
3	EDO	A	502	-	-	1/1/1/1	-
3	EDO	B	502	-	-	0/1/1/1	-
3	EDO	C	503	-	-	0/1/1/1	-
2	TRS	C	501	-	-	0/9/9/9	-
2	TRS	A	501	-	-	0/9/9/9	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	502	EDO	O1-C1-C2-O2
2	B	501	TRS	C2-C-C3-O3
3	A	502	EDO	O1-C1-C2-O2
2	B	501	TRS	N-C-C3-O3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	501	TRS	1	0
2	A	501	TRS	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	444/444 (100%)	0.01	9 (2%) 65 70	8, 19, 34, 52	1 (0%)
1	B	443/444 (99%)	-0.03	4 (0%) 81 86	8, 19, 34, 46	4 (0%)
1	C	441/444 (99%)	0.18	16 (3%) 46 50	10, 20, 41, 61	0
1	D	443/444 (99%)	0.27	19 (4%) 40 43	7, 21, 44, 62	2 (0%)
All	All	1771/1776 (99%)	0.11	48 (2%) 56 61	7, 20, 38, 62	7 (0%)

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	306	LEU	4.5
1	C	305	GLY	3.5
1	D	305	GLY	3.4
1	C	309	ALA	3.4
1	C	272	ASP	3.3

### 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 oligosaccharides 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
2	TRS	C	501	8/8	0.90	0.10	17,22,22,25	0
3	EDO	A	502	4/4	0.92	0.09	20,22,22,23	0
3	EDO	C	503	4/4	0.92	0.09	21,22,22,24	0
3	EDO	D	503	4/4	0.93	0.11	17,20,23,27	0
3	EDO	C	502	4/4	0.94	0.10	17,22,28,33	0
2	TRS	D	502	8/8	0.94	0.07	21,23,24,25	0
2	TRS	A	501	8/8	0.94	0.08	15,16,22,22	0
3	EDO	B	502	4/4	0.96	0.06	16,18,20,21	0
2	TRS	B	501	8/8	0.96	0.06	11,15,18,22	0
4	NA	B	504	1/1	0.96	0.10	26,26,26,26	0
4	NA	D	505	1/1	0.96	0.07	25,25,25,25	0
4	NA	B	505	1/1	0.97	0.07	28,28,28,28	0
4	NA	C	505	1/1	0.97	0.09	25,25,25,25	0
4	NA	D	504	1/1	0.97	0.08	24,24,24,24	0
4	NA	A	504	1/1	0.97	0.04	25,25,25,25	0
4	NA	D	501	1/1	0.98	0.04	21,21,21,21	0
4	NA	C	504	1/1	0.98	0.06	24,24,24,24	0
4	NA	A	505	1/1	0.98	0.06	25,25,25,25	0
4	NA	A	506	1/1	0.99	0.05	20,20,20,20	0
4	NA	B	503	1/1	0.99	0.06	20,20,20,20	0
4	NA	A	503	1/1	0.99	0.08	20,20,20,20	0

## 6.5 Other polymers

There are no such residues in this entry.