



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 15, 2026 – 07:18 PM EDT

PDB ID : 9MZ5 / pdb\_00009mz5  
Title : EatA-EatI complex from Mycobacterium avium subsp. paratuberculosis  
Authors : Benedict, S.T.; Moynihan, P.J.  
Deposited on : 2025-01-22  
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	:	2022.3.0, CSD as543be (2022)
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

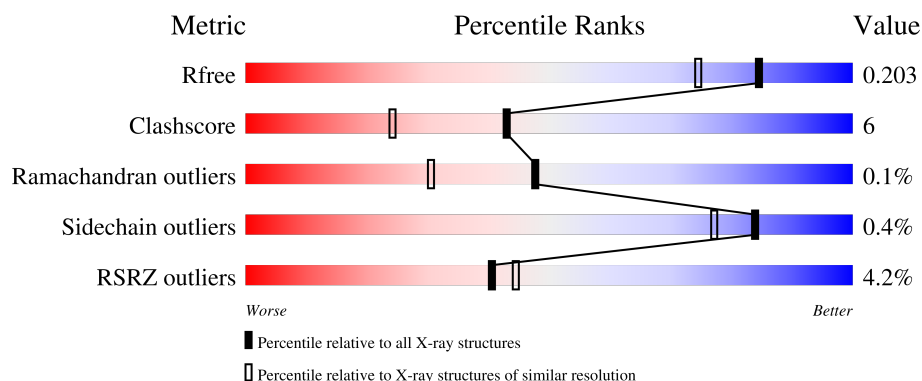
# 1 Overall quality at a glance

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	338	<div> <div>2%</div> <div>91%</div> <div>7%</div> <div>.</div> </div>
1	B	338	<div> <div>5%</div> <div>86%</div> <div>10%</div> <div>.</div> </div>
2	C	160	<div> <div>4%</div> <div>74%</div> <div>8%</div> <div>17%</div> </div>
2	D	160	<div> <div>6%</div> <div>82%</div> <div>9%</div> <div>9%</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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NO3	A	402	-	-	X	-
4	EDO	D	202	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 14691 atoms, of which 6649 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 EatA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	B	324	Total	C	H	N	O	S	0	4	0
			4732	1530	2304	417	476	5			
1	A	333	Total	C	H	N	O	S	0	7	0
			4886	1581	2377	435	489	4			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	MET	-	initiating methionine	UNP Q743L4
B	0	ALA	-	cloning artifact	UNP Q743L4
B	328	GLY	-	expression tag	UNP Q743L4
B	329	SER	-	expression tag	UNP Q743L4
B	330	GLY	-	expression tag	UNP Q743L4
B	331	HIS	-	expression tag	UNP Q743L4
B	332	HIS	-	expression tag	UNP Q743L4
B	333	HIS	-	expression tag	UNP Q743L4
B	334	HIS	-	expression tag	UNP Q743L4
B	335	HIS	-	expression tag	UNP Q743L4
B	336	HIS	-	expression tag	UNP Q743L4
A	-1	MET	-	initiating methionine	UNP Q743L4
A	0	ALA	-	cloning artifact	UNP Q743L4
A	328	GLY	-	expression tag	UNP Q743L4
A	329	SER	-	expression tag	UNP Q743L4
A	330	GLY	-	expression tag	UNP Q743L4
A	331	HIS	-	expression tag	UNP Q743L4
A	332	HIS	-	expression tag	UNP Q743L4
A	333	HIS	-	expression tag	UNP Q743L4
A	334	HIS	-	expression tag	UNP Q743L4
A	335	HIS	-	expression tag	UNP Q743L4
A	336	HIS	-	expression tag	UNP Q743L4

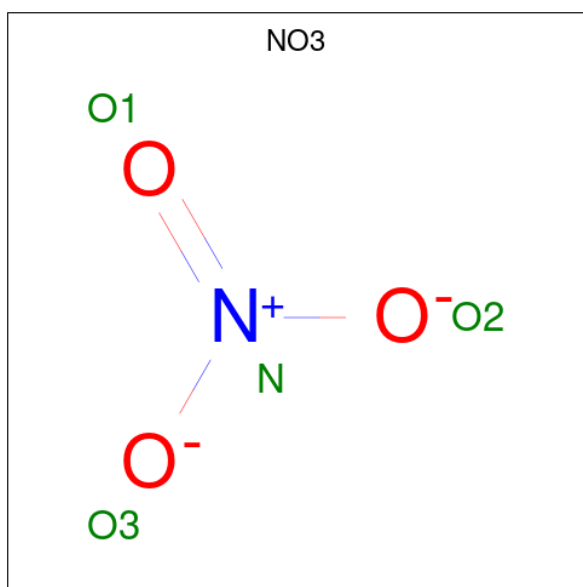
- Molecule 2 is a protein called EatI from Mycobacterium avium subsp. paratuberculosis.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	133	Total	C	H	N	O	S	0	1	0
			1916	612	940	166	193	5			
2	D	145	Total	C	H	N	O	S	0	2	0
			2069	659	1016	179	210	5			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-9	MET	-	initiating methionine	UNP Q743L5
C	-8	ALA	-	expression tag	UNP Q743L5
C	-7	HIS	-	expression tag	UNP Q743L5
C	-6	HIS	-	expression tag	UNP Q743L5
C	-5	HIS	-	expression tag	UNP Q743L5
C	-4	HIS	-	expression tag	UNP Q743L5
C	-3	HIS	-	expression tag	UNP Q743L5
C	-2	HIS	-	expression tag	UNP Q743L5
C	-1	GLY	-	expression tag	UNP Q743L5
C	0	SER	-	expression tag	UNP Q743L5
C	1	GLY	-	expression tag	UNP Q743L5
D	-9	MET	-	initiating methionine	UNP Q743L5
D	-8	ALA	-	expression tag	UNP Q743L5
D	-7	HIS	-	expression tag	UNP Q743L5
D	-6	HIS	-	expression tag	UNP Q743L5
D	-5	HIS	-	expression tag	UNP Q743L5
D	-4	HIS	-	expression tag	UNP Q743L5
D	-3	HIS	-	expression tag	UNP Q743L5
D	-2	HIS	-	expression tag	UNP Q743L5
D	-1	GLY	-	expression tag	UNP Q743L5
D	0	SER	-	expression tag	UNP Q743L5
D	1	GLY	-	expression tag	UNP Q743L5

- Molecule 3 is NITRATE ION (CCD ID: NO3) (formula: NO<sub>3</sub>).



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

- Molecule 4 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
4	A	1	Total	C	H	O	0	0
			10	2	6	2		
4	D	1	Total	C	H	O	0	0
			10	2	6	2		

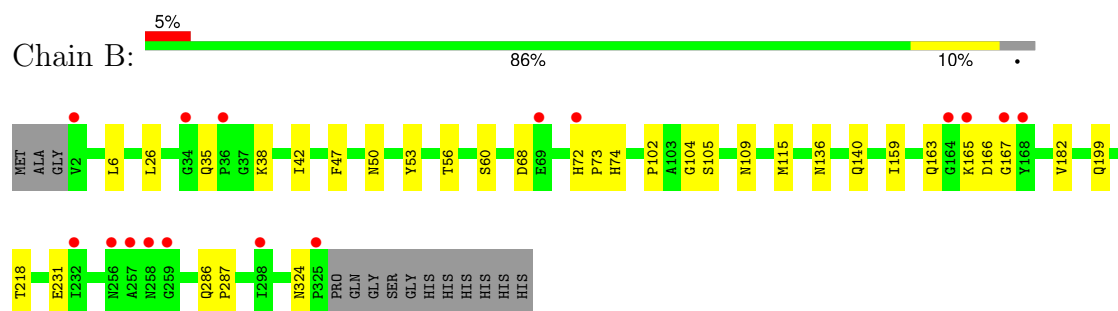
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	313	Total	O	0	0
			313	313		
5	C	146	Total	O	0	0
			146	146		
5	A	393	Total	O	0	2
			395	395		
5	D	190	Total	O	0	0
			190	190		

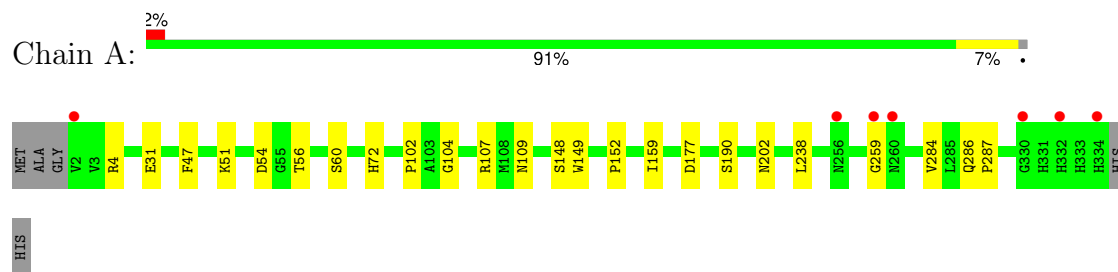
### 3 Residue-property plots [i](#)

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.

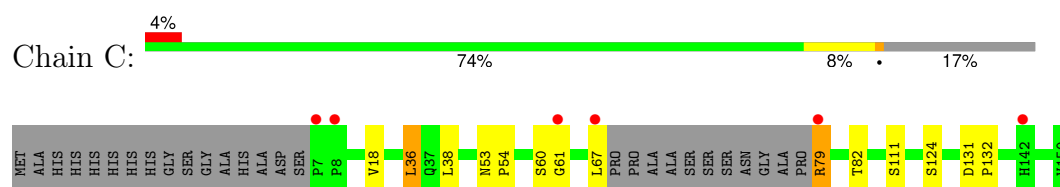
- Molecule 1: EatA



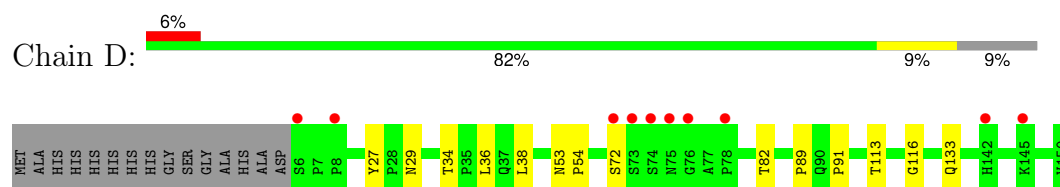
- Molecule 1: EatA



- Molecule 2: EatI from *Mycobacterium avium* subsp. *paratuberculosis*



- Molecule 2: EatI from *Mycobacterium avium* subsp. *paratuberculosis*



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.26Å 101.47Å 118.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	59.13 – 1.65 59.13 – 1.65	Depositor EDS
% Data completeness (in resolution range)	69.8 (59.13-1.65) 69.8 (59.13-1.65)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.61 (at 1.65Å)	Xtriage
Refinement program	PHENIX 1.21.2_5419	Depositor
R, $R_{free}$	0.163 , 0.204 0.163 , 0.203	Depositor DCC
$R_{free}$ test set	3510 reflections (3.59%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.2	Xtriage
Anisotropy	0.044	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 42.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	14691	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	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 23.31 % 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 4.7838e-03. 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: NO3, 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	0.48	0/2608	0.64	0/3581
1	B	0.49	1/2512 (0.0%)	0.62	0/3445
2	C	0.49	0/1005	0.66	0/1380
2	D	0.48	0/1089	0.65	0/1500
All	All	0.49	1/7214 (0.0%)	0.64	0/9906

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	73	PRO	N-CD	5.78	1.55	1.47

There are no bond angle outliers.

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	2509	2377	2376	21	0
1	B	2428	2304	2303	29	0
2	C	976	940	940	17	0
2	D	1053	1016	1016	12	0
3	A	4	0	0	6	0
3	B	8	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	4	0	0	0	0
3	D	8	0	0	0	0
4	A	4	6	6	1	0
4	D	4	6	6	4	0
5	A	395	0	0	13	10
5	B	313	0	0	15	8
5	C	146	0	0	11	1
5	D	190	0	0	4	7
All	All	8042	6649	6647	84	14

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 6.

The worst 5 of 84 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)
3:A:402:NO3:O3	5:A:501:HOH:O	1.60	1.20
2:C:67:LEU:O	5:C:302:HOH:O	1.67	1.12
1:B:136:ASN:ND2	5:B:504:HOH:O	1.94	0.99
1:B:166:ASP:O	5:B:501:HOH:O	1.83	0.95
1:B:167:GLY:HA3	5:B:501:HOH:O	1.68	0.94

The worst 5 of 14 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:B:751:HOH:O	5:A:773:HOH:O[4_445]	1.31	0.89
5:B:731:HOH:O	5:A:804:HOH:O[4_545]	1.35	0.85
5:A:800:HOH:O	5:D:442:HOH:O[1_455]	1.45	0.75
5:A:737:HOH:O	5:A:802:HOH:O[4_535]	1.46	0.74
5:B:724:HOH:O	5:B:765:HOH:O[4_545]	1.79	0.41

## 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	338/338 (100%)	325 (96%)	13 (4%)	0	100	100
1	B	326/338 (96%)	318 (98%)	7 (2%)	1 (0%)	36	21
2	C	129/160 (81%)	125 (97%)	4 (3%)	0	100	100
2	D	144/160 (90%)	142 (99%)	2 (1%)	0	100	100
All	All	937/996 (94%)	910 (97%)	26 (3%)	1 (0%)	48	30

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	26	LEU

### 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	264/260 (102%)	264 (100%)	0	100	100
1	B	254/260 (98%)	254 (100%)	0	100	100
2	C	113/130 (87%)	111 (98%)	2 (2%)	51	30
2	D	122/130 (94%)	121 (99%)	1 (1%)	73	60
All	All	753/780 (96%)	750 (100%)	3 (0%)	84	77

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	36	LEU
2	C	79	ARG
2	D	72	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	140	GLN
1	A	267	ASN
2	D	37	GLN
1	A	332	HIS
2	C	53	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

8 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	EDO	D	202	-	3,3,3	0.40	0	2,2,2	0.23	0
3	NO3	C	201	-	1,3,3	0.18	0	0,3,3	-	-
3	NO3	B	402	-	1,3,3	0.30	0	0,3,3	-	-
4	EDO	A	401	-	3,3,3	0.29	0	2,2,2	0.44	0
3	NO3	D	203	-	1,3,3	0.18	0	0,3,3	-	-
3	NO3	B	401	-	1,3,3	0.11	0	0,3,3	-	-
3	NO3	D	201	-	1,3,3	0.19	0	0,3,3	-	-
3	NO3	A	402	-	1,3,3	0.10	0	0,3,3	-	-

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
4	EDO	D	202	-	-	1/1/1/1	-
4	EDO	A	401	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	202	EDO	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	202	EDO	4	0
4	A	401	EDO	1	0
3	B	401	NO3	1	0
3	A	402	NO3	6	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	333/338 (98%)	-0.30	7 (2%) 63 69	5, 16, 36, 56	7 (2%)
1	B	324/338 (95%)	0.08	16 (4%) 35 39	7, 20, 48, 68	4 (1%)
2	C	133/160 (83%)	-0.01	6 (4%) 38 42	10, 21, 41, 54	1 (0%)
2	D	145/160 (90%)	-0.21	10 (6%) 23 26	8, 18, 48, 71	2 (1%)
All	All	935/996 (93%)	-0.11	39 (4%) 40 44	5, 18, 43, 71	14 (1%)

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	76	GLY	3.9
1	B	72	HIS	3.8
1	B	325	PRO	3.7
2	C	67	LEU	3.6
2	D	75	ASN	3.6

### 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
3	NO3	B	401	4/4	0.79	0.11	29,30,33,34	4
4	EDO	A	401	4/4	0.81	0.13	29,36,48,58	0
3	NO3	A	402	4/4	0.84	0.11	17,21,22,27	4
3	NO3	D	203	4/4	0.85	0.10	35,41,43,46	0
3	NO3	B	402	4/4	0.86	0.10	17,19,30,32	4
4	EDO	D	202	4/4	0.88	0.15	20,27,33,37	0
3	NO3	C	201	4/4	0.93	0.08	24,30,33,35	0
3	NO3	D	201	4/4	0.94	0.07	19,24,26,28	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.