



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 02:34 am BST

PDB ID : 6V0Q  
Title : Crystal structure of the bromodomain of human BRD7 bound to TG003  
Authors : Karim, M.R.; Chan, A.; Schonbrunn, E.  
Deposited on : 2019-11-19  
Resolution : 1.69 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

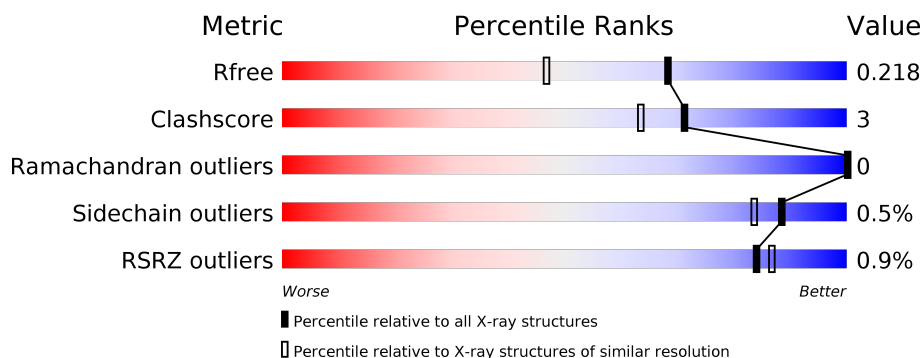
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.69 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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

Mol	Chain	Length	Quality of chain
1	A	123	<div> <div>%</div> <div> <div></div> <div>89%</div> <div>7%</div> <div>•</div> </div> </div>
1	B	123	<div> <div></div> <div>89%</div> <div>6%</div> <div>5%</div> </div>
1	C	123	<div> <div>%</div> <div> <div></div> <div>87%</div> <div>8%</div> <div>5%</div> </div> </div>
1	D	123	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>8%</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 4503 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 Bromodomain-containing protein 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	119	Total	C	N	O	S	0	3	0
			988	633	161	184	10			
1	B	117	Total	C	N	O	S	0	4	0
			977	628	159	180	10			
1	C	117	Total	C	N	O	S	0	4	0
			977	627	159	180	11			
1	D	117	Total	C	N	O	S	0	5	0
			978	627	159	182	10			

There are 8 discrepancies between the modelled and reference sequences:

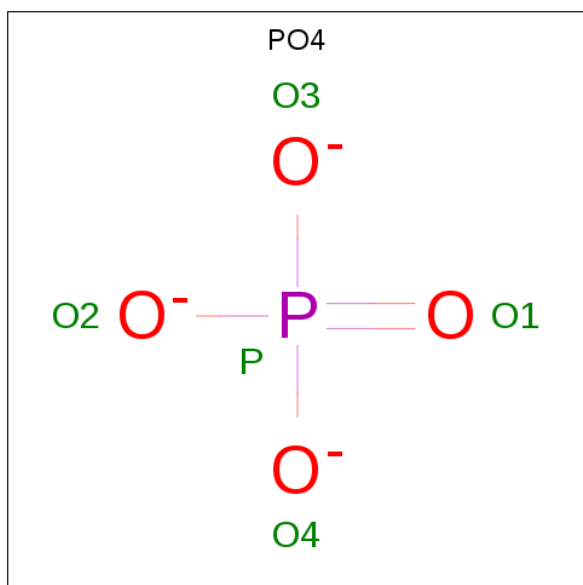
Chain	Residue	Modelled	Actual	Comment	Reference
A	128	GLU	-	expression tag	UNP Q9NPI1
A	129	SER	-	expression tag	UNP Q9NPI1
B	128	GLU	-	expression tag	UNP Q9NPI1
B	129	SER	-	expression tag	UNP Q9NPI1
C	128	GLU	-	expression tag	UNP Q9NPI1
C	129	SER	-	expression tag	UNP Q9NPI1
D	128	GLU	-	expression tag	UNP Q9NPI1
D	129	SER	-	expression tag	UNP Q9NPI1

- Molecule 2 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
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).

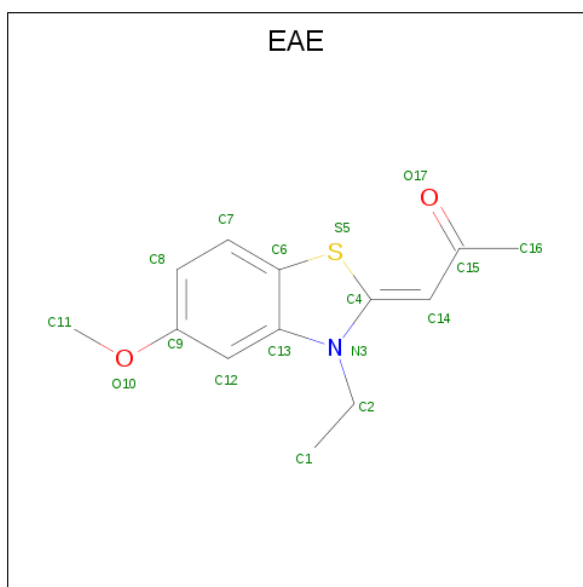


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Cl	0	0
			1	1		

- Molecule 5 is (1 {Z})-1-(3-ethyl-5-methoxy-1,3-benzothiazol-2-ylidene)propan-2-one (three-letter code: EAE) (formula: C<sub>13</sub>H<sub>15</sub>NO<sub>2</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	S	0	0
			17	13	1	2	1		
5	B	1	Total	C	N	O	S	0	0
			17	13	1	2	1		
5	C	1	Total	C	N	O	S	0	0
			17	13	1	2	1		
5	D	1	Total	C	N	O	S	0	0
			17	13	1	2	1		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			6	3	3		

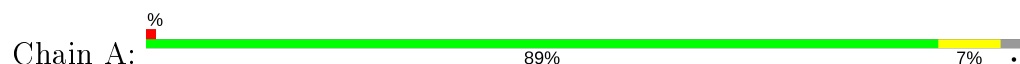
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	142	Total	O	0	0
			142	142		
7	B	138	Total	O	0	0
			138	138		
7	C	92	Total	O	0	0
			92	92		
7	D	115	Total	O	0	0
			115	115		

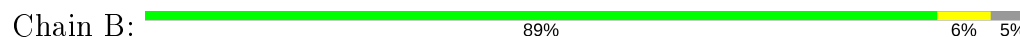
### 3 Residue-property plots [i](#)

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

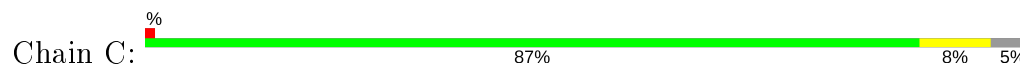
- Molecule 1: Bromodomain-containing protein 7



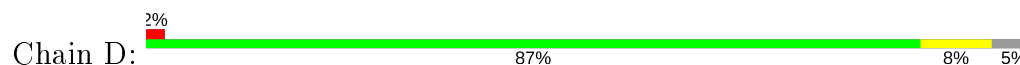
- Molecule 1: Bromodomain-containing protein 7



- Molecule 1: Bromodomain-containing protein 7



- Molecule 1: Bromodomain-containing protein 7



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.97Å 36.78Å 118.96Å 90.00° 101.77° 90.00°	Depositor
Resolution (Å)	37.68 – 1.69 58.71 – 1.69	Depositor EDS
% Data completeness (in resolution range)	98.5 (37.68-1.69) 98.5 (58.71-1.69)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.10 (at 1.69Å)	Xtriage
Refinement program	PHENIX 1.12	Depositor
R, $R_{free}$	0.181 , 0.218 0.182 , 0.218	Depositor DCC
$R_{free}$ test set	1193 reflections (2.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.7	Xtriage
Anisotropy	0.242	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 50.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.006 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4503	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	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 34.86 % 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 6.3822e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, PO4, EAE, EDO, 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	0.40	0/1013	0.50	0/1357
1	B	0.39	0/1005	0.50	0/1346
1	C	0.38	0/1005	0.46	0/1345
1	D	0.36	0/1009	0.48	0/1351
All	All	0.38	0/4032	0.48	0/5399

There are no bond length outliers.

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	988	0	1005	10	0
1	B	977	0	1001	5	0
1	C	977	0	999	6	0
1	D	978	0	1000	6	0
2	A	8	0	12	2	0
2	C	4	0	6	1	0
2	D	4	0	6	0	0
3	A	5	0	0	0	0
4	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	17	0	0	0	0
5	B	17	0	0	1	0
5	C	17	0	0	0	0
5	D	17	0	0	0	0
6	A	6	0	8	3	0
7	A	142	0	0	1	0
7	B	138	0	0	0	0
7	C	92	0	0	2	0
7	D	115	0	0	0	0
All	All	4503	0	4037	28	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 3.

The worst 5 of 28 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:150:LYS:NZ	6:A:306:GOL:H2	1.93	0.82
1:D:162:ASP:OD2	1:D:169:SER:HB2	1.79	0.81
1:C:176:MET:HG3	1:C:203[B]:MET:HG3	1.85	0.58
1:C:141:ASN:O	1:C:145:ARG:HG3	2.07	0.53
1:A:217:TYR:CE2	2:A:302:EDO:H12	2.44	0.53

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	120/123 (98%)	120 (100%)	0	0	100	100
1	B	119/123 (97%)	119 (100%)	0	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	119/123 (97%)	119 (100%)	0	0	100	100
1	D	120/123 (98%)	119 (99%)	1 (1%)	0	100	100
All	All	478/492 (97%)	477 (100%)	1 (0%)	0	100	100

There are no Ramachandran outliers to report.

### 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	113/114 (99%)	113 (100%)	0	100	100
1	B	112/114 (98%)	112 (100%)	0	100	100
1	C	112/114 (98%)	111 (99%)	1 (1%)	78	70
1	D	113/114 (99%)	112 (99%)	1 (1%)	78	70
All	All	450/456 (99%)	448 (100%)	2 (0%)	88	87

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	201	LYS
1	D	132	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	191	GLN

### 5.3.3 RNA ⓘ

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 11 ligands modelled in this entry, 1 is monoatomic - leaving 10 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	EAE	B	301	-	15,18,18	1.35	2 (13%)	13,25,25	2.21	4 (30%)
2	EDO	C	301	-	3,3,3	0.44	0	2,2,2	0.37	0
5	EAE	A	305	-	15,18,18	1.34	2 (13%)	13,25,25	2.40	4 (30%)
5	EAE	D	302	-	15,18,18	1.29	2 (13%)	13,25,25	3.07	4 (30%)
2	EDO	A	301	-	3,3,3	0.46	0	2,2,2	0.37	0
2	EDO	A	302	-	3,3,3	0.31	0	2,2,2	0.65	0
6	GOL	A	306	-	5,5,5	1.03	0	5,5,5	1.31	0
5	EAE	C	302	-	15,18,18	1.31	2 (13%)	13,25,25	3.01	3 (23%)
3	PO4	A	303	-	4,4,4	0.85	0	6,6,6	0.46	0
2	EDO	D	301	-	3,3,3	0.47	0	2,2,2	0.30	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
5	EAE	B	301	-	-	0/6/8/8	0/2/2/2
2	EDO	C	301	-	-	0/1/1/1	-
5	EAE	A	305	-	-	1/6/8/8	0/2/2/2
5	EAE	D	302	-	-	4/6/8/8	0/2/2/2

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EDO	A	301	-	-	0/1/1/1	-
2	EDO	A	302	-	-	1/1/1/1	-
5	EAE	C	302	-	-	3/6/8/8	0/2/2/2
6	GOL	A	306	-	-	2/4/4/4	-
2	EDO	D	301	-	-	1/1/1/1	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	302	EAE	C12-C9	2.98	1.42	1.37
5	B	301	EAE	C12-C9	2.96	1.42	1.37
5	C	302	EAE	C12-C9	2.95	1.42	1.37
5	A	305	EAE	C7-C8	2.63	1.42	1.36
5	A	305	EAE	C12-C9	2.47	1.41	1.37

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	302	EAE	C1-C2-N3	7.11	123.50	111.49
5	D	302	EAE	C1-C2-N3	6.70	122.81	111.49
5	D	302	EAE	C12-C13-C6	-5.91	115.59	120.56
5	C	302	EAE	C12-C13-C6	-5.48	115.96	120.56
5	B	301	EAE	C1-C2-N3	5.46	120.72	111.49

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	D	302	EAE	C1-C2-N3-C13
5	D	302	EAE	C1-C2-N3-C4
6	A	306	GOL	O1-C1-C2-O2
5	C	302	EAE	C1-C2-N3-C4
5	C	302	EAE	C12-C9-O10-C11

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	301	EAE	1	0
2	C	301	EDO	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	302	EDO	2	0
6	A	306	GOL	3	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	119/123 (96%)	-0.25	1 (0%) 86 88	12, 22, 42, 58	0
1	B	117/123 (95%)	-0.26	0 100 100	13, 24, 43, 66	0
1	C	117/123 (95%)	-0.02	1 (0%) 84 87	16, 30, 53, 63	0
1	D	117/123 (95%)	-0.15	2 (1%) 70 74	14, 26, 48, 68	0
All	All	470/492 (95%)	-0.17	4 (0%) 84 87	12, 26, 47, 68	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	170	MET	3.1
1	D	170	MET	2.6
1	D	169	SER	2.5
1	A	132	GLU	2.1

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

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	GOL	A	306	6/6	0.49	0.19	52,57,62,63	0
2	EDO	D	301	4/4	0.70	0.14	51,55,55,55	0
2	EDO	C	301	4/4	0.82	0.15	60,61,61,62	0
5	EAE	C	302	17/17	0.88	0.13	22,31,57,58	0
5	EAE	D	302	17/17	0.90	0.13	22,31,48,49	0
4	CL	A	304	1/1	0.90	0.08	50,50,50,50	0
2	EDO	A	301	4/4	0.91	0.12	32,35,40,43	0
2	EDO	A	302	4/4	0.93	0.25	24,44,47,49	0
3	PO4	A	303	5/5	0.95	0.20	60,62,63,66	0
5	EAE	B	301	17/17	0.98	0.08	11,14,20,30	0
5	EAE	A	305	17/17	0.98	0.07	10,13,17,21	0

## 6.5 Other polymers

There are no such residues in this entry.