



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 20, 2022 – 12:52 pm GMT

PDB ID : 7PWJ  
Title : Human dUTPase in complex.  
Authors : Sanz-Frasquet, C.; Marina, A.  
Deposited on : 2021-10-06  
Resolution : 1.94 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.3
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)	:	2.31.3

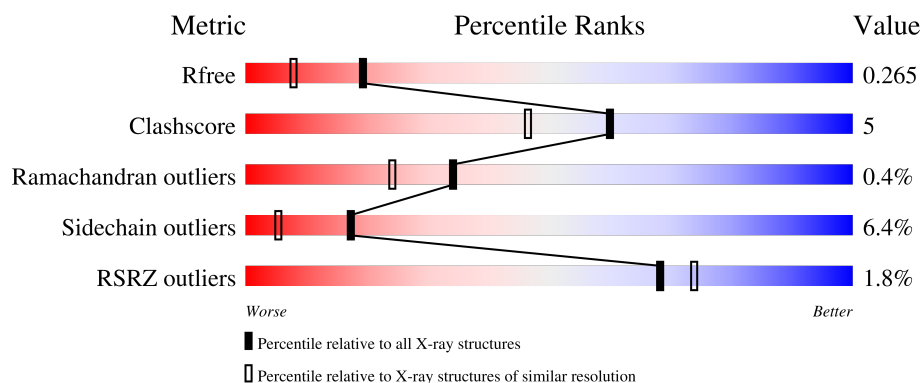
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.94 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	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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	CCC	156	<div> <div>4%</div> <div>80%12%6%</div> </div>
1	EEE	156	<div> <div>85%8%8%</div> </div>
1	FFF	156	<div> <div>%84%10%5%</div> </div>
2	AAA	145	<div> <div>%81%13%6%</div> </div>
2	BBB	145	<div> <div>%86%9%5%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	DDD	145	

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	EDO	FFF	201	-	-	X	-
5	GOL	AAA	207	-	-	X	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7052 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 Orf20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	CCC	146	Total	C	N	O	S	0	1	0
			1119	725	176	216	2			
1	EEE	144	Total	C	N	O	S	0	0	0
			1158	749	183	224	2			
1	FFF	148	Total	C	N	O	S	0	1	0
			1194	769	189	234	2			

- Molecule 2 is a protein called Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	AAA	137	Total	C	N	O	S	0	1	0
			1045	665	179	197	4			
2	BBB	138	Total	C	N	O	S	0	0	0
			1034	660	178	193	3			
2	DDD	138	Total	C	N	O	S	0	1	0
			1041	665	179	193	4			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	-3	TYR	-	expression tag	UNP P33316
AAA	-2	PHE	-	expression tag	UNP P33316
AAA	-1	GLN	-	expression tag	UNP P33316
BBB	-3	TYR	-	expression tag	UNP P33316
BBB	-2	PHE	-	expression tag	UNP P33316
BBB	-1	GLN	-	expression tag	UNP P33316
DDD	-3	TYR	-	expression tag	UNP P33316
DDD	-2	PHE	-	expression tag	UNP P33316
DDD	-1	GLN	-	expression tag	UNP P33316

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



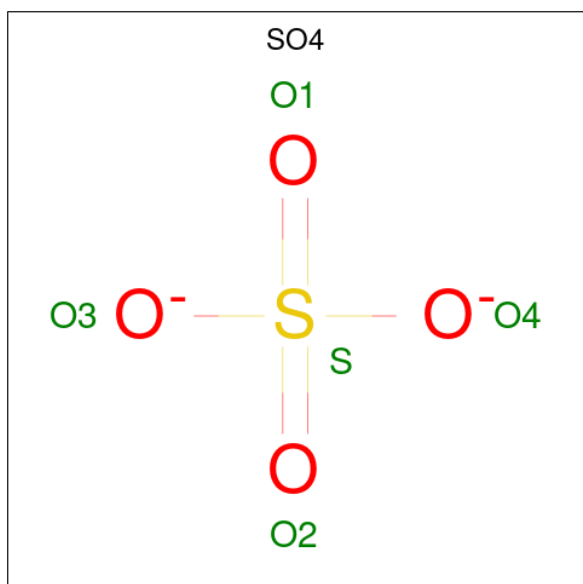
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	CCC	1	Total	C	O	0	0
			4	2	2		
3	CCC	1	Total	C	O	0	0
			4	2	2		
3	CCC	1	Total	C	O	0	0
			4	2	2		
3	AAA	1	Total	C	O	0	0
			4	2	2		
3	AAA	1	Total	C	O	0	0
			4	2	2		
3	AAA	1	Total	C	O	0	0
			4	2	2		
3	BBB	1	Total	C	O	0	0
			4	2	2		
3	BBB	1	Total	C	O	0	0
			4	2	2		
3	BBB	1	Total	C	O	0	0
			4	2	2		
3	DDD	1	Total	C	O	0	0
			4	2	2		
3	DDD	1	Total	C	O	0	0
			4	2	2		
3	DDD	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	EEE	1	Total	C	O	0	0
			4	2	2		
3	EEE	1	Total	C	O	0	0
			4	2	2		
3	EEE	1	Total	C	O	0	0
			4	2	2		
3	FFF	1	Total	C	O	0	0
			4	2	2		
3	FFF	1	Total	C	O	0	0
			4	2	2		
3	FFF	1	Total	C	O	0	0
			4	2	2		

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



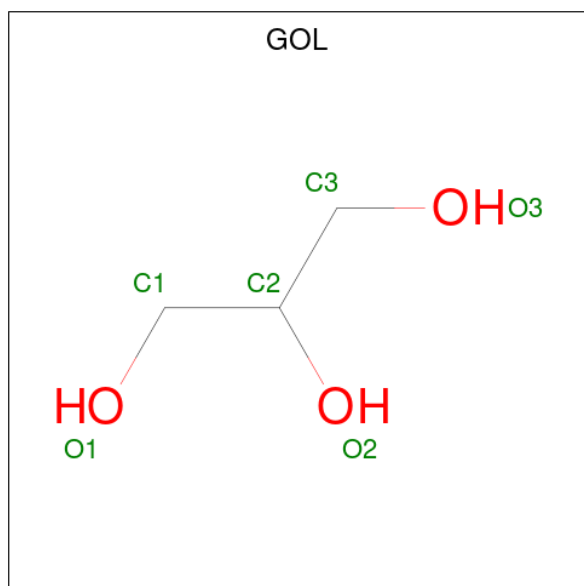
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	AAA	1	Total	O	S	0	0
			5	4	1		
4	AAA	1	Total	O	S	0	0
			5	4	1		
4	BBB	1	Total	O	S	0	0
			5	4	1		
4	DDD	1	Total	O	S	0	0
			5	4	1		
4	EEE	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	FFF	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	AAA	1	Total	C	O	0	0
			6	3	3		
5	BBB	1	Total	C	O	0	0
			6	3	3		
5	BBB	1	Total	C	O	0	0
			6	3	3		
5	DDD	1	Total	C	O	0	0
			6	3	3		
5	FFF	1	Total	C	O	0	0
			6	3	3		
5	FFF	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	CCC	22	Total	O	0	0
			22	22		
6	AAA	63	Total	O	0	0
			63	63		

*Continued on next page...*

*Continued from previous page...*


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	BBB	55	Total 55	O 55	0	0
6	DDD	76	Total 76	O 76	0	0
6	EEE	49	Total 49	O 49	0	0
6	FFF	50	Total 50	O 50	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: Orf20

Chain CCC: 




- Molecule 1: Orf20

Chain EEE: 




- Molecule 1: Orf20

Chain FFF: 



- Molecule 2: Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial

Chain AAA: 

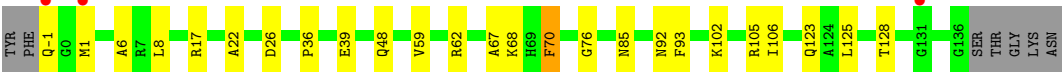
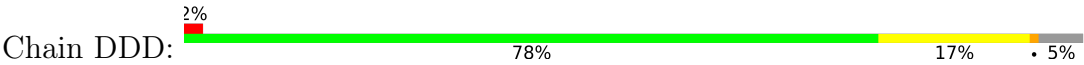


- Molecule 2: Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial

Chain BBB: 



- Molecule 2: Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.72Å 81.80Å 198.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	99.20 – 1.94 99.20 – 1.94	Depositor EDS
% Data completeness (in resolution range)	98.5 (99.20-1.94) 98.5 (99.20-1.94)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.40 (at 1.94Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.214 , 0.262 0.220 , 0.265	Depositor DCC
$R_{free}$ test set	4665 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.6	Xtriage
Anisotropy	0.432	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 51.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7052	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.88% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, GOL, SO4

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	CCC	0.77	1/1147 (0.1%)	0.80	0/1556
1	EEE	0.70	0/1183	0.79	0/1598
1	FFF	0.70	0/1219	0.77	0/1647
2	AAA	0.70	0/1066	0.88	0/1439
2	BBB	0.65	0/1054	0.86	0/1421
2	DDD	0.77	1/1062 (0.1%)	0.87	1/1433 (0.1%)
All	All	0.72	2/6731 (0.0%)	0.83	1/9094 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	DDD	-1	GLN	C-O	7.04	1.36	1.23
1	CCC	51	ARG	C-O	5.63	1.34	1.23

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	DDD	62	ARG	NE-CZ-NH2	6.30	123.45	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts ⓘ

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	CCC	1119	0	1002	9	0
1	EEE	1158	0	1095	5	0
1	FFF	1194	0	1126	15	0
2	AAA	1045	0	1005	18	0
2	BBB	1034	0	997	4	0
2	DDD	1041	0	1009	12	0
3	AAA	16	0	24	4	0
3	BBB	16	0	24	2	0
3	CCC	12	0	18	1	0
3	DDD	12	0	18	1	0
3	EEE	12	0	18	0	0
3	FFF	12	0	18	5	0
4	AAA	10	0	0	0	0
4	BBB	5	0	0	0	0
4	DDD	5	0	0	0	0
4	EEE	5	0	0	0	0
4	FFF	5	0	0	0	0
5	AAA	6	0	8	6	0
5	BBB	12	0	16	0	0
5	DDD	6	0	8	3	0
5	FFF	12	0	16	2	0
6	AAA	63	0	0	3	0
6	BBB	55	0	0	0	0
6	CCC	22	0	0	0	0
6	DDD	76	0	0	0	0
6	EEE	49	0	0	0	0
6	FFF	50	0	0	0	0
All	All	7052	0	6402	61	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 5.

All (61) 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)
2:DDD:85:ASN:HD22	5:DDD:202:GOL:H32	1.40	0.85
2:AAA:121:GLU:H	1:FFF:56:ASN:HD21	1.21	0.85
5:AAA:207:GOL:H12	3:FFF:201:EDO:H21	1.69	0.75
1:FFF:13:HIS:HA	5:FFF:204:GOL:H2	1.72	0.70
2:DDD:26:ASP:HB3	2:DDD:105:ARG:HD3	1.80	0.64
1:FFF:74[B]:ARG:HH11	1:FFF:74[B]:ARG:HG3	1.61	0.64
2:AAA:59:VAL:HG23	2:AAA:78:ILE:CD1	2.29	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:CCC:97:MET:CE	1:CCC:125:LEU:HB3	2.30	0.61
2:AAA:121:GLU:N	1:FFF:56:ASN:HD21	1.96	0.60
2:BBB:62:ARG:HA	3:BBB:202:EDO:H12	1.83	0.59
2:BBB:95:LYS:H	3:BBB:204:EDO:H21	1.66	0.59
5:AAA:207:GOL:H12	3:FFF:201:EDO:C2	2.32	0.58
2:DDD:70:PHE:CD1	2:DDD:93:PHE:HB2	2.39	0.58
1:FFF:145:LEU:HG	1:FFF:149:LYS:HE3	1.86	0.56
1:FFF:14:TYR:O	1:FFF:18:ILE:HD12	2.05	0.56
1:FFF:74[B]:ARG:HG3	1:FFF:74[B]:ARG:NH1	2.21	0.55
1:EEE:12:THR:OG1	1:EEE:14:TYR:CD2	2.56	0.55
1:CCC:97:MET:HE2	1:CCC:125:LEU:HB3	1.92	0.52
2:DDD:6:ALA:HB3	2:DDD:48:GLN:HE21	1.76	0.51
1:CCC:97:MET:HE3	1:CCC:125:LEU:HB3	1.93	0.51
2:AAA:123:GLN:H	3:AAA:205:EDO:H21	1.76	0.51
1:CCC:59:GLU:HA	1:CCC:69:SER:OG	2.11	0.50
2:AAA:0:GLY:HA2	3:AAA:206:EDO:H12	1.92	0.50
1:FFF:13:HIS:CA	5:FFF:204:GOL:H2	2.39	0.49
2:AAA:48:GLN:NE2	2:BBB:125:LEU:H	2.10	0.49
1:EEE:72:LEU:N	1:EEE:72:LEU:HD23	2.28	0.49
2:AAA:62:ARG:HD2	6:AAA:337:HOH:O	2.12	0.49
2:AAA:122:VAL:HG21	2:AAA:125:LEU:HD23	1.96	0.48
1:EEE:110:ASP:HB2	1:EEE:149:LYS:HE2	1.94	0.48
2:AAA:59:VAL:HG23	2:AAA:78:ILE:HD13	1.96	0.47
2:AAA:59:VAL:HB	2:AAA:76:GLY:HA2	1.97	0.47
2:AAA:85:ASN:HD22	5:AAA:207:GOL:C3	2.28	0.47
5:AAA:207:GOL:C3	3:FFF:201:EDO:H21	2.44	0.47
1:CCC:140:ASP:HB2	3:CCC:203:EDO:O1	2.15	0.47
2:DDD:8:LEU:CD2	3:DDD:204:EDO:H21	2.45	0.47
6:AAA:311:HOH:O	1:FFF:55:VAL:HG13	2.15	0.46
2:DDD:68:LYS:HD2	1:FFF:114:SER:HB2	1.98	0.46
2:DDD:36:PRO:HB2	2:DDD:39:GLU:HG3	1.98	0.45
2:AAA:59:VAL:HG12	2:AAA:73:VAL:HG11	1.98	0.45
2:AAA:121:GLU:H	1:FFF:56:ASN:ND2	2.03	0.45
3:AAA:203:EDO:H22	3:AAA:206:EDO:H22	1.99	0.45
2:BBB:78:ILE:CD1	2:BBB:86:VAL:HG13	2.48	0.44
2:DDD:85:ASN:HD22	5:DDD:202:GOL:C3	2.22	0.44
2:DDD:85:ASN:HB2	5:DDD:202:GOL:H31	1.98	0.44
1:CCC:94:PHE:HA	1:CCC:126:LEU:HD11	2.00	0.43
2:AAA:0:GLY:HA2	3:AAA:206:EDO:C1	2.49	0.43
2:AAA:121:GLU:OE1	1:FFF:55:VAL:HG13	2.18	0.43
2:DDD:17:ARG:CG	2:DDD:22:ALA:HB3	2.49	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:AAA:11:HIS:HD2	6:AAA:355:HOH:O	2.02	0.43
1:EEE:103:LYS:HE3	1:EEE:141:TYR:CZ	2.55	0.42
1:EEE:12:THR:HG23	1:EEE:14:TYR:H	1.84	0.42
2:AAA:85:ASN:HD22	5:AAA:207:GOL:H32	1.84	0.42
5:AAA:207:GOL:C1	3:FFF:201:EDO:H21	2.45	0.42
1:FFF:52:ASN:HB3	3:FFF:206:EDO:H21	2.01	0.42
2:AAA:40:LYS:HE2	2:DDD:67:ALA:O	2.19	0.42
1:FFF:125:LEU:HD23	1:FFF:151:ILE:HD12	2.02	0.41
1:CCC:100:TYR:CD2	1:CCC:144:VAL:HG21	2.56	0.41
1:CCC:32:LEU:O	1:CCC:36:THR:HG23	2.21	0.40
1:CCC:129:LEU:HD23	1:CCC:129:LEU:HA	1.93	0.40
2:DDD:59:VAL:HB	2:DDD:76:GLY:HA2	2.03	0.40
1:FFF:70:TYR:O	1:FFF:74[A]:ARG:HG2	2.22	0.40

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	CCC	145/156 (93%)	133 (92%)	11 (8%)	1 (1%)	22	11
1	EEE	142/156 (91%)	138 (97%)	4 (3%)	0	100	100
1	FFF	147/156 (94%)	141 (96%)	5 (3%)	1 (1%)	22	11
2	AAA	136/145 (94%)	134 (98%)	2 (2%)	0	100	100
2	BBB	134/145 (92%)	131 (98%)	2 (2%)	1 (1%)	22	11
2	DDD	137/145 (94%)	135 (98%)	2 (2%)	0	100	100
All	All	841/903 (93%)	812 (97%)	26 (3%)	3 (0%)	34	24

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	BBB	132	SER
1	FFF	9	GLU
1	CCC	50	ASN

### 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	CCC	105/141 (74%)	96 (91%)	9 (9%)	10	2
1	EEE	118/141 (84%)	113 (96%)	5 (4%)	30	14
1	FFF	124/141 (88%)	119 (96%)	5 (4%)	31	16
2	AAA	100/113 (88%)	93 (93%)	7 (7%)	15	4
2	BBB	97/113 (86%)	90 (93%)	7 (7%)	14	3
2	DDD	100/113 (88%)	91 (91%)	9 (9%)	9	1
All	All	644/762 (84%)	602 (94%)	42 (6%)	17	5

All (42) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	CCC	10	LEU
1	CCC	14	TYR
1	CCC	19	LYS
1	CCC	32	LEU
1	CCC	96	LYS
1	CCC	97	MET
1	CCC	123	ILE
1	CCC	136	VAL
1	CCC	148	TYR
2	AAA	19	SER
2	AAA	40	LYS
2	AAA	51	LEU
2	AAA	70	PHE
2	AAA	102	LYS
2	AAA	106	ILE
2	AAA	130	ARG

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	BBB	9	SER
2	BBB	19	SER
2	BBB	55	CYS
2	BBB	65	LEU
2	BBB	70	PHE
2	BBB	119	ILE
2	BBB	122	VAL
2	DDD	1	MET
2	DDD	70	PHE
2	DDD	92	ASN
2	DDD	102[A]	LYS
2	DDD	102[B]	LYS
2	DDD	106	ILE
2	DDD	123	GLN
2	DDD	125	LEU
2	DDD	128	THR
1	EEE	52	ASN
1	EEE	88	LEU
1	EEE	90	ASP
1	EEE	128	LEU
1	EEE	148	TYR
1	FFF	10	LEU
1	FFF	13	HIS
1	FFF	55	VAL
1	FFF	128	LEU
1	FFF	148	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

32 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	$\# Z  > 2$
3	EDO	CCC	202	-	3,3,3	0.10	0	2,2,2	0.20	0
3	EDO	AAA	203	-	3,3,3	0.26	0	2,2,2	0.33	0
3	EDO	AAA	205	-	3,3,3	0.12	0	2,2,2	0.14	0
5	GOL	FFF	203	-	5,5,5	0.17	0	5,5,5	0.41	0
4	SO4	DDD	201	-	4,4,4	0.43	0	6,6,6	0.12	0
3	EDO	DDD	205	-	3,3,3	0.14	0	2,2,2	0.23	0
3	EDO	BBB	205	-	3,3,3	0.28	0	2,2,2	0.55	0
3	EDO	AAA	204	-	3,3,3	0.21	0	2,2,2	0.44	0
3	EDO	BBB	204	-	3,3,3	0.21	0	2,2,2	0.41	0
5	GOL	AAA	207	-	5,5,5	0.20	0	5,5,5	0.59	0
4	SO4	BBB	207	-	4,4,4	0.39	0	6,6,6	0.34	0
3	EDO	FFF	206	-	3,3,3	0.10	0	2,2,2	0.31	0
5	GOL	DDD	202	-	5,5,5	0.14	0	5,5,5	0.37	0
4	SO4	AAA	201	-	4,4,4	0.33	0	6,6,6	0.16	0
3	EDO	EEE	201	-	3,3,3	0.13	0	2,2,2	0.31	0
5	GOL	FFF	204	-	5,5,5	0.14	0	5,5,5	0.52	0
3	EDO	EEE	204	-	3,3,3	0.29	0	2,2,2	0.12	0
5	GOL	BBB	201	-	5,5,5	0.18	0	5,5,5	0.53	0
4	SO4	FFF	202	-	4,4,4	0.34	0	6,6,6	0.07	0
3	EDO	BBB	202	-	3,3,3	0.25	0	2,2,2	0.47	0
3	EDO	EEE	203	-	3,3,3	0.34	0	2,2,2	0.39	0
4	SO4	AAA	202	-	4,4,4	0.30	0	6,6,6	0.09	0
3	EDO	CCC	203	-	3,3,3	0.14	0	2,2,2	0.06	0
5	GOL	BBB	203	-	5,5,5	0.29	0	5,5,5	0.80	0
4	SO4	EEE	202	-	4,4,4	0.36	0	6,6,6	0.06	0
3	EDO	FFF	201	-	3,3,3	0.25	0	2,2,2	0.74	0
3	EDO	AAA	206	-	3,3,3	0.04	0	2,2,2	0.47	0
3	EDO	DDD	203	-	3,3,3	0.25	0	2,2,2	0.26	0
3	EDO	CCC	201	-	3,3,3	0.12	0	2,2,2	0.14	0
3	EDO	FFF	205	-	3,3,3	0.12	0	2,2,2	0.06	0
3	EDO	DDD	204	-	3,3,3	0.22	0	2,2,2	0.12	0
3	EDO	BBB	206	-	3,3,3	0.18	0	2,2,2	0.43	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
3	EDO	CCC	202	-	-	0/1/1/1	-
3	EDO	AAA	203	-	-	0/1/1/1	-
3	EDO	AAA	205	-	-	1/1/1/1	-
5	GOL	FFF	203	-	-	4/4/4/4	-
3	EDO	DDD	205	-	-	0/1/1/1	-
3	EDO	BBB	205	-	-	0/1/1/1	-
3	EDO	AAA	204	-	-	0/1/1/1	-
3	EDO	BBB	204	-	-	1/1/1/1	-
5	GOL	AAA	207	-	-	2/4/4/4	-
3	EDO	FFF	206	-	-	0/1/1/1	-
5	GOL	DDD	202	-	-	4/4/4/4	-
3	EDO	EEE	201	-	-	0/1/1/1	-
5	GOL	FFF	204	-	-	0/4/4/4	-
3	EDO	EEE	204	-	-	0/1/1/1	-
5	GOL	BBB	201	-	-	2/4/4/4	-
3	EDO	BBB	202	-	-	0/1/1/1	-
3	EDO	EEE	203	-	-	0/1/1/1	-
3	EDO	CCC	203	-	-	0/1/1/1	-
5	GOL	BBB	203	-	-	4/4/4/4	-
3	EDO	FFF	201	-	-	1/1/1/1	-
3	EDO	AAA	206	-	-	1/1/1/1	-
3	EDO	DDD	203	-	-	1/1/1/1	-
3	EDO	CCC	201	-	-	1/1/1/1	-
3	EDO	FFF	205	-	-	1/1/1/1	-
3	EDO	DDD	204	-	-	0/1/1/1	-
3	EDO	BBB	206	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (23) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	DDD	202	GOL	C1-C2-C3-O3
5	FFF	203	GOL	O1-C1-C2-C3
5	FFF	203	GOL	C1-C2-C3-O3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
5	AAA	207	GOL	O2-C2-C3-O3
5	AAA	207	GOL	C1-C2-C3-O3
5	BBB	201	GOL	O1-C1-C2-C3
5	BBB	203	GOL	O1-C1-C2-C3
5	BBB	203	GOL	C1-C2-C3-O3
5	DDD	202	GOL	O1-C1-C2-C3
5	BBB	201	GOL	O1-C1-C2-O2
5	DDD	202	GOL	O2-C2-C3-O3
5	FFF	203	GOL	O2-C2-C3-O3
3	AAA	206	EDO	O1-C1-C2-O2
5	BBB	203	GOL	O1-C1-C2-O2
5	FFF	203	GOL	O1-C1-C2-O2
5	BBB	203	GOL	O2-C2-C3-O3
3	BBB	204	EDO	O1-C1-C2-O2
3	DDD	203	EDO	O1-C1-C2-O2
3	FFF	201	EDO	O1-C1-C2-O2
3	AAA	205	EDO	O1-C1-C2-O2
3	CCC	201	EDO	O1-C1-C2-O2
5	DDD	202	GOL	O1-C1-C2-O2
3	FFF	205	EDO	O1-C1-C2-O2

There are no ring outliers.

12 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	AAA	203	EDO	1	0
3	AAA	205	EDO	1	0
3	BBB	204	EDO	1	0
5	AAA	207	GOL	6	0
3	FFF	206	EDO	1	0
5	DDD	202	GOL	3	0
5	FFF	204	GOL	2	0
3	BBB	202	EDO	1	0
3	CCC	203	EDO	1	0
3	FFF	201	EDO	4	0
3	AAA	206	EDO	3	0
3	DDD	204	EDO	1	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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<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	CCC	146/156 (93%)	0.50	7 (4%) 30 38	42, 63, 92, 103	0
1	EEE	144/156 (92%)	0.24	0 100 100	37, 51, 73, 98	0
1	FFF	148/156 (94%)	0.27	1 (0%) 87 91	36, 52, 78, 98	1 (0%)
2	AAA	137/145 (94%)	0.34	2 (1%) 73 79	31, 48, 78, 89	0
2	BBB	138/145 (95%)	0.33	2 (1%) 75 80	35, 49, 86, 99	0
2	DDD	138/145 (95%)	0.30	3 (2%) 62 69	33, 44, 78, 84	1 (0%)
All	All	851/903 (94%)	0.33	15 (1%) 68 74	31, 51, 83, 103	2 (0%)

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	FFF	8	ALA	6.4
1	CCC	50	ASN	4.5
1	CCC	133	LYS	3.3
2	BBB	131	GLY	3.2
1	CCC	51	ARG	3.1
2	DDD	1	MET	2.7
2	DDD	131	GLY	2.7
1	CCC	38	PHE	2.5
1	CCC	134	ILE	2.5
1	CCC	39	SER	2.3
2	DDD	-1	GLN	2.2
2	AAA	131	GLY	2.2
2	BBB	129	GLU	2.2
2	AAA	132	SER	2.1
1	CCC	100	TYR	2.0

## 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 monosaccharides 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	EDO	EEE	201	4/4	0.53	0.36	93,95,98,100	0
3	EDO	EEE	203	4/4	0.67	0.19	61,74,76,79	0
3	EDO	DDD	205	4/4	0.70	0.16	84,85,90,96	0
3	EDO	AAA	206	4/4	0.71	0.21	89,91,92,94	0
5	GOL	FFF	204	6/6	0.71	0.18	57,63,65,65	0
3	EDO	CCC	202	4/4	0.73	0.15	91,92,92,92	0
3	EDO	CCC	201	4/4	0.75	0.15	76,78,82,82	0
5	GOL	FFF	203	6/6	0.79	0.12	83,87,90,91	0
3	EDO	DDD	203	4/4	0.81	0.23	64,65,70,70	0
3	EDO	AAA	204	4/4	0.82	0.21	72,75,80,81	0
3	EDO	BBB	205	4/4	0.83	0.21	54,65,66,67	0
4	SO4	AAA	202	5/5	0.83	0.18	97,105,111,122	0
3	EDO	AAA	205	4/4	0.84	0.13	64,65,66,69	0
3	EDO	AAA	203	4/4	0.84	0.14	65,67,71,73	0
3	EDO	FFF	205	4/4	0.86	0.13	75,75,76,80	0
3	EDO	CCC	203	4/4	0.88	0.14	74,75,75,77	0
4	SO4	FFF	202	5/5	0.88	0.16	118,127,131,132	0
4	SO4	EEE	202	5/5	0.90	0.13	111,112,120,125	0
3	EDO	FFF	206	4/4	0.90	0.10	73,76,78,80	0
5	GOL	BBB	203	6/6	0.91	0.18	49,58,65,67	0
3	EDO	BBB	204	4/4	0.91	0.13	70,72,75,77	0
5	GOL	BBB	201	6/6	0.91	0.15	63,70,73,74	0
3	EDO	DDD	204	4/4	0.92	0.28	54,58,60,63	0
5	GOL	DDD	202	6/6	0.92	0.15	48,57,64,65	0
5	GOL	AAA	207	6/6	0.93	0.17	47,56,66,68	0
3	EDO	FFF	201	4/4	0.93	0.24	43,61,62,72	0
3	EDO	BBB	202	4/4	0.93	0.39	48,66,69,71	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	EDO	BBB	206	4/4	0.94	0.10	67,72,75,77	0
3	EDO	EEE	204	4/4	0.94	0.19	67,68,69,85	0
4	SO4	AAA	201	5/5	0.98	0.14	59,61,65,72	0
4	SO4	BBB	207	5/5	0.98	0.12	50,52,58,60	0
4	SO4	DDD	201	5/5	0.99	0.12	40,43,48,55	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.