



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 20, 2022 – 03:28 pm GMT

PDB ID : 7PWX  
Title : Structure of Mycobacterium tuberculosis dUTPase at 2.75Å.  
Authors : Marina, A.; Sanz-Frasquet, C.  
Deposited on : 2021-10-07  
Resolution : 2.75 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

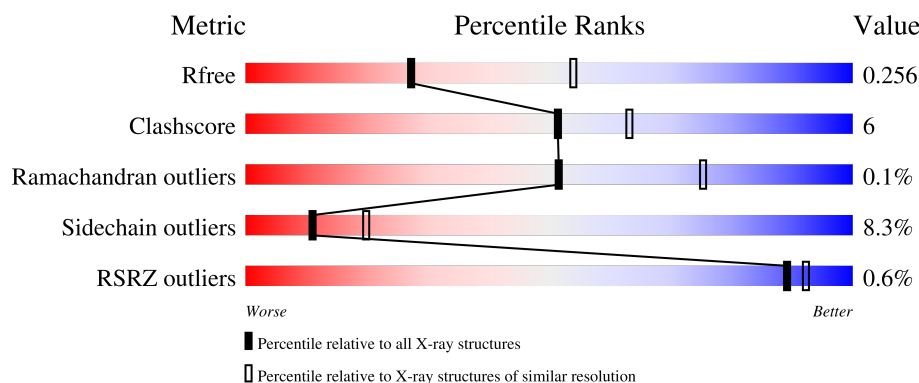
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	HHH	136	
1	III	136	
1	JJJ	136	
1	KKK	136	
1	LLL	136	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	MMM	136	<div><div></div><div>77%</div><div>17%</div><div></div><div></div></div>
2	AAA	147	<div><div>4%</div><div></div><div>72%</div><div>7%</div><div></div><div>20%</div></div>
2	CCC	147	<div><div></div><div>82%</div><div>16%</div><div></div><div></div></div>
2	DDD	147	<div><div>%</div><div></div><div>84%</div><div>14%</div><div></div><div></div></div>
2	EEE	147	<div><div>%</div><div></div><div>89%</div><div>10%</div><div></div><div></div></div>
2	FFF	147	<div><div>%</div><div></div><div>81%</div><div>17%</div><div></div><div></div></div>
2	GGG	147	<div><div></div><div>71%</div><div>14%</div><div></div><div></div><div>13%</div></div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12287 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 Deoxyuridine 5'-triphosphate nucleotidohydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	HHH	134	Total	C	N	O	S	0	0	0
			946	602	167	176	1			
1	III	131	Total	C	N	O	S	0	0	0
			950	606	167	176	1			
1	JJJ	131	Total	C	N	O	S	0	0	0
			952	604	168	179	1			
1	KKK	135	Total	C	N	O	S	0	0	0
			946	604	172	169	1			
1	LLL	136	Total	C	N	O	S	0	0	0
			977	620	178	177	2			
1	MMM	131	Total	C	N	O	S	0	0	0
			950	603	172	174	1			

- Molecule 2 is a protein called Orf20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	CCC	146	Total	C	N	O	S	0	0	0
			1174	757	186	229	2			
2	DDD	145	Total	C	N	O	S	0	0	0
			1167	757	180	228	2			
2	EEE	146	Total	C	N	O	S	0	0	0
			1148	742	178	226	2			
2	FFF	146	Total	C	N	O	S	0	2	0
			1184	768	186	228	2			
2	GGG	128	Total	C	N	O	S	0	0	0
			919	588	145	184	2			
2	AAA	118	Total	C	N	O	S	0	0	0
			787	499	130	156	2			

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



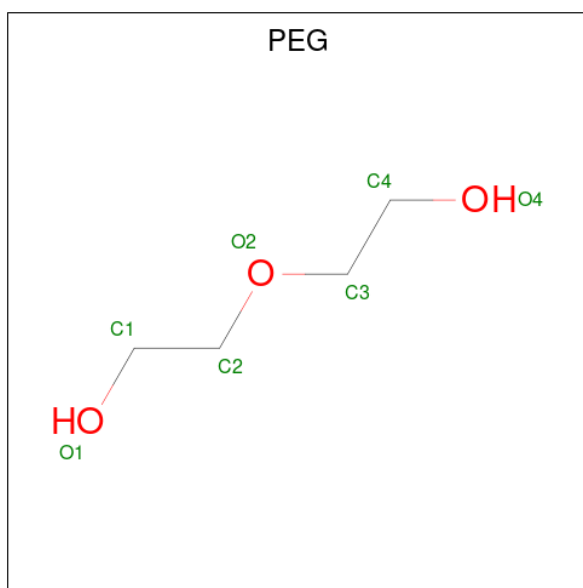
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	HHH	1	Total	C	N	O	0	0
			8	4	1	3		
3	JJJ	1	Total	C	N	O	0	0
			8	4	1	3		
3	LLL	1	Total	C	N	O	0	0
			8	4	1	3		
3	LLL	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 4 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
4	III	1	Total C O 4 2 2	0	0
4	JJJ	1	Total C O 4 2 2	0	0
4	JJJ	1	Total C O 4 2 2	0	0
4	JJJ	1	Total C O 4 2 2	0	0
4	DDD	1	Total C O 4 2 2	0	0
4	DDD	1	Total C O 4 2 2	0	0
4	FFF	1	Total C O 4 2 2	0	0
4	GGG	1	Total C O 4 2 2	0	0

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	JJJ	1	Total C O 7 4 3	0	0
5	KKK	1	Total C O 7 4 3	0	0
5	MMM	1	Total C O 7 4 3	0	0
5	FFF	1	Total C O 7 4 3	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	HHH	7	Total O 7 7	0	0
6	III	8	Total O 8 8	0	0
6	JJJ	9	Total O 9 9	0	0
6	KKK	2	Total O 2 2	0	0
6	LLL	8	Total O 8 8	0	0
6	MMM	7	Total O 7 7	0	0
6	CCC	7	Total O 7 7	0	0
6	DDD	12	Total O 12 12	0	0
6	EEE	4	Total O 4 4	0	0
6	FFF	17	Total O 17 17	0	0
6	GGG	8	Total O 8 8	0	0
6	AAA	6	Total O 6 6	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: Deoxyuridine 5'-triphosphate nucleotidohydrolase

Chain HHH:  85% 12% ..




- Molecule 1: Deoxyuridine 5'-triphosphate nucleotidohydrolase

Chain III:  74% 20% . .




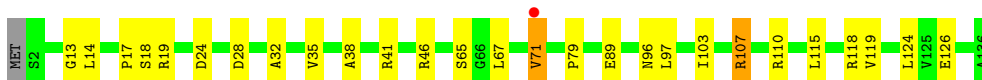
- Molecule 1: Deoxyuridine 5'-triphosphate nucleotidohydrolase

Chain JJJ:  83% 12% . .




- Molecule 1: Deoxyuridine 5'-triphosphate nucleotidohydrolase

Chain KKK:  79% 18% ..




- Molecule 1: Deoxyuridine 5'-triphosphate nucleotidohydrolase

Chain LLL:  87% 12% .



- Molecule 1: Deoxyuridine 5'-triphosphate nucleotidohydrolase

Chain MMM:  77% 17% . .





- Molecule 2: Orf20

Chain CCC: 82% 16% ..



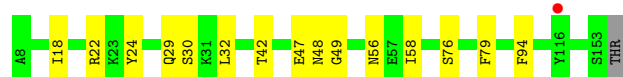
- Molecule 2: Orf20

Chain DDD: 84% 14% ..



- Molecule 2: Orf20

Chain EEE: 89% 10% .



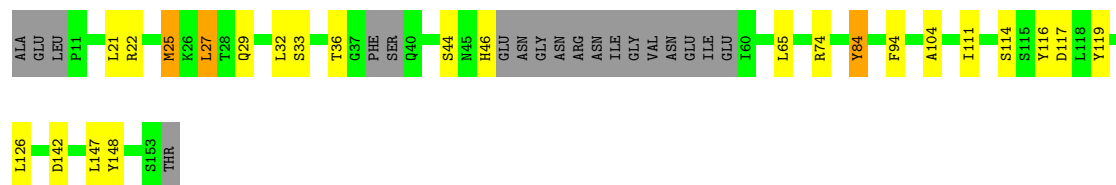
- Molecule 2: Orf20

Chain FFF: 81% 17% ..



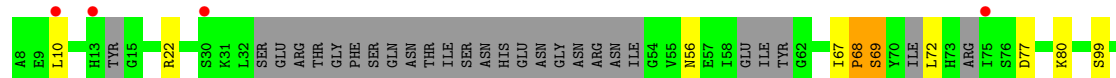
- Molecule 2: Orf20

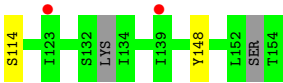
Chain GGG: 71% 14% . 13%



- Molecule 2: Orf20

Chain AAA: 72% 7% . 20%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.41Å 150.43Å 107.37Å 90.00° 98.85° 90.00°	Depositor
Resolution (Å)	75.33 – 2.75 75.21 – 2.75	Depositor EDS
% Data completeness (in resolution range)	95.1 (75.33-2.75) 95.1 (75.21-2.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.88 (at 2.73Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.201 , 0.259 0.203 , 0.256	Depositor DCC
$R_{free}$ test set	3102 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	63.7	Xtriage
Anisotropy	0.113	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 40.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	12287	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% 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: PEG, 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	HHH	0.70	0/960	0.83	0/1311
1	III	0.65	0/964	0.86	0/1314
1	JJJ	0.65	0/967	0.85	0/1322
1	KKK	0.69	0/960	0.86	0/1312
1	LLL	0.68	0/992	0.87	0/1354
1	MMM	0.68	0/965	0.87	0/1319
2	AAA	0.78	0/795	0.82	0/1076
2	CCC	0.68	0/1199	0.78	0/1622
2	DDD	0.68	0/1192	0.78	0/1611
2	EEE	0.69	0/1173	0.76	0/1590
2	FFF	0.67	0/1214	0.84	2/1640 (0.1%)
2	GGG	0.79	0/937	0.81	0/1273
All	All	0.69	0/12318	0.83	2/16744 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	FFF	48[A]	ASN	CB-CA-C	-8.13	94.14	110.40
2	FFF	48[B]	ASN	CB-CA-C	-8.13	94.14	110.40

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	HHH	946	0	934	13	0
1	III	950	0	964	17	0
1	JJJ	952	0	968	13	0
1	KKK	946	0	960	14	0
1	LLL	977	0	1002	17	0
1	MMM	950	0	979	18	0
2	AAA	787	0	572	3	0
2	CCC	1174	0	1112	13	0
2	DDD	1167	0	1088	12	0
2	EEE	1148	0	1046	7	0
2	FFF	1184	0	1119	12	0
2	GGG	919	0	735	15	0
3	HHH	8	0	12	5	0
3	JJJ	8	0	12	0	0
3	LLL	16	0	24	5	0
4	DDD	8	0	12	0	0
4	FFF	4	0	6	0	0
4	GGG	4	0	6	0	0
4	III	4	0	6	0	0
4	JJJ	12	0	18	0	0
5	FFF	7	0	10	0	0
5	JJJ	7	0	10	0	0
5	KKK	7	0	10	0	0
5	MMM	7	0	10	2	0
6	AAA	6	0	0	0	0
6	CCC	7	0	0	0	0
6	DDD	12	0	0	0	0
6	EEE	4	0	0	0	0
6	FFF	17	0	0	0	0
6	GGG	8	0	0	1	0
6	HHH	7	0	0	0	0
6	III	8	0	0	0	0
6	JJJ	9	0	0	0	0
6	KKK	2	0	0	0	0
6	LLL	8	0	0	0	0
6	MMM	7	0	0	0	0
All	All	12287	0	11615	134	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 6.

All (134) 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:LLL:122:VAL:HG12	1:MMM:3:THR:HG22	1.30	1.08
1:LLL:122:VAL:CG1	1:MMM:3:THR:HG22	1.99	0.92
2:FFF:10:LEU:HD21	2:FFF:55:VAL:HA	1.61	0.82
2:GGG:27:LEU:HD13	2:GGG:27:LEU:O	1.87	0.74
3:LLL:202:TRS:H12	1:MMM:60:LEU:HD13	1.70	0.73
1:HHH:44:LEU:HD22	1:HHH:91:LYS:HB3	1.72	0.71
3:HHH:201:TRS:O3	1:JJJ:117:GLN:NE2	2.24	0.69
1:LLL:117:GLN:OE1	3:LLL:202:TRS:N	2.27	0.67
2:GGG:29:GLN:O	2:GGG:33:SER:N	2.29	0.66
2:FFF:10:LEU:CD2	2:FFF:55:VAL:HA	2.26	0.65
1:MMM:46:ARG:HG2	1:MMM:89:GLU:OE2	1.98	0.63
1:LLL:104:VAL:HG13	1:LLL:104:VAL:O	1.98	0.63
2:DDD:139:ILE:CG1	2:DDD:143:TYR:HD2	2.12	0.62
2:CCC:22:ARG:HD3	2:CCC:32:LEU:HD22	1.80	0.62
2:DDD:22:ARG:HD3	2:DDD:32:LEU:HD22	1.81	0.61
2:DDD:139:ILE:CG1	2:DDD:143:TYR:CD2	2.84	0.60
1:LLL:1:MET:C	1:LLL:1:MET:SD	2.80	0.59
1:MMM:46:ARG:HG2	1:MMM:89:GLU:CD	2.23	0.59
2:GGG:25:MET:HB3	2:GGG:27:LEU:HD12	1.84	0.58
1:KKK:28:ASP:HB2	1:KKK:110:ARG:NH1	2.18	0.57
1:HHH:126:GLU:H	2:EEE:56:ASN:HD21	1.53	0.56
1:KKK:38:ALA:HB3	1:KKK:41:ARG:HG3	1.87	0.56
2:FFF:94:PHE:HA	2:FFF:126:LEU:HD11	1.88	0.56
1:KKK:17:PRO:HB3	1:KKK:28:ASP:O	2.07	0.55
2:FFF:35:ARG:HD2	2:FFF:64:GLY:O	2.07	0.54
1:MMM:60:LEU:CD1	1:MMM:117:GLN:HE21	2.19	0.54
2:GGG:84:TYR:H	2:GGG:84:TYR:HD1	1.55	0.54
1:III:77:ASN:ND2	1:JJJ:65:SER:HB3	2.23	0.54
1:JJJ:15:PRO:HG2	1:JJJ:30:TYR:CE2	2.43	0.54
1:LLL:9:ARG:NH1	1:LLL:14:LEU:O	2.39	0.54
1:HHH:62:HIS:CE1	1:JJJ:79:PRO:HG2	2.43	0.53
1:III:120:GLU:HG3	1:JJJ:57:MET:SD	2.49	0.53
1:MMM:7:ILE:HD11	1:MMM:49:VAL:CG1	2.39	0.53
1:KKK:19:ARG:NH1	1:KKK:24:ASP:O	2.42	0.52
3:HHH:201:TRS:C3	1:III:117:GLN:HE22	2.22	0.52
1:MMM:28:ASP:OD2	1:MMM:110:ARG:NH1	2.42	0.52
2:CCC:45:ASN:HA	2:CCC:48:ASN:ND2	2.24	0.52
1:KKK:96:ASN:ND2	1:KKK:103:ILE:HD12	2.25	0.51
3:HHH:201:TRS:H22	1:III:79:PRO:HG3	1.91	0.51
2:CCC:48:ASN:HD22	2:CCC:49:GLY:N	2.08	0.51
3:HHH:201:TRS:H32	1:III:117:GLN:NE2	2.25	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:III:125:VAL:HG13	1:JJJ:4:THR:HG23	1.92	0.51
2:CCC:18:ILE:HD11	2:CCC:58:ILE:HD13	1.92	0.51
2:GGG:27:LEU:HD21	2:GGG:32:LEU:CB	2.41	0.51
1:HHH:44:LEU:CD2	1:HHH:91:LYS:HB3	2.40	0.51
1:LLL:122:VAL:HA	1:MMM:3:THR:O	2.11	0.51
1:MMM:45:VAL:HG21	1:MMM:94:LEU:HD12	1.93	0.51
1:III:67:LEU:HG	1:III:73:LEU:HD23	1.92	0.50
1:LLL:79:PRO:HG2	1:MMM:62:HIS:CE1	2.46	0.50
1:LLL:104:VAL:O	1:LLL:104:VAL:CG1	2.60	0.50
2:DDD:69:SER:O	2:DDD:72:LEU:HB2	2.12	0.50
2:GGG:74:ARG:HB3	2:GGG:119:TYR:CE2	2.47	0.50
2:EEE:29:GLN:NE2	2:EEE:47:GLU:OE1	2.45	0.49
1:KKK:124:LEU:HD23	1:LLL:7:ILE:HD13	1.95	0.49
2:CCC:70:TYR:OH	2:CCC:117:ASP:OD2	2.21	0.49
3:LLL:202:TRS:HN2	1:MMM:117:GLN:HE22	1.60	0.48
1:KKK:79:PRO:HG3	3:LLL:202:TRS:H31	1.95	0.48
2:FFF:110:ASP:HB2	2:FFF:149:LYS:HE2	1.96	0.48
2:GGG:44:SER:C	2:GGG:46:HIS:H	2.17	0.48
2:CCC:18:ILE:HD11	2:CCC:58:ILE:CD1	2.44	0.48
2:GGG:74:ARG:NH2	2:GGG:117:ASP:OD2	2.44	0.48
3:HHH:201:TRS:O2	1:JJJ:60:LEU:HD13	2.14	0.47
1:III:67:LEU:HB3	1:III:73:LEU:HB3	1.97	0.47
1:JJJ:14:LEU:HD11	1:JJJ:32:ALA:HA	1.95	0.47
1:LLL:7:ILE:CG2	1:LLL:16:LEU:HD11	2.45	0.47
1:LLL:126:GLU:H	2:CCC:56:ASN:HD21	1.63	0.47
2:DDD:104:ALA:HB1	2:DDD:111:ILE:HG13	1.96	0.47
1:LLL:70:ARG:HH12	2:DDD:152:LEU:HA	1.80	0.47
2:CCC:45:ASN:HA	2:CCC:48:ASN:HD21	1.80	0.46
1:MMM:64:ARG:NH1	5:MMM:201:PEG:O4	2.44	0.46
2:FFF:100:TYR:HB3	2:FFF:144:VAL:HG11	1.98	0.46
2:GGG:22:ARG:HA	2:GGG:27:LEU:HD11	1.97	0.46
2:AAA:67:ILE:O	2:AAA:68:PRO:C	2.54	0.46
2:DDD:101:VAL:HG12	2:DDD:118:LEU:HG	1.97	0.45
2:DDD:136:VAL:O	2:DDD:139:ILE:HG22	2.16	0.45
2:FFF:82:LYS:HB3	2:FFF:82:LYS:HE3	1.72	0.45
1:KKK:14:LEU:HD11	1:KKK:32:ALA:HA	1.99	0.45
2:DDD:48:ASN:HD22	2:DDD:49:GLY:N	2.15	0.45
2:CCC:10:LEU:CD2	2:CCC:58:ILE:CG1	2.95	0.45
1:HHH:130:PHE:O	1:HHH:134:GLY:O	2.35	0.45
2:GGG:65:LEU:HA	6:GGG:305:HOH:O	2.17	0.45
1:HHH:76:VAL:CG1	1:III:68:ALA:HB1	2.47	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:CCC:148:TYR:CE1	2:CCC:152:LEU:HD11	2.51	0.44
1:MMM:51:VAL:HG22	1:MMM:90:ILE:HD11	1.98	0.44
2:FFF:94:PHE:CE1	2:FFF:122:THR:HG21	2.53	0.44
1:KKK:126:GLU:HA	1:LLL:7:ILE:O	2.18	0.44
2:CCC:48:ASN:HD22	2:CCC:48:ASN:C	2.21	0.44
2:DDD:17:ILE:HG22	2:DDD:21:LEU:HD12	2.00	0.44
1:HHH:68:ALA:HB1	1:JJJ:76:VAL:HG12	1.99	0.43
1:KKK:71:VAL:HG23	1:KKK:103:ILE:HG12	2.00	0.43
2:EEE:22:ARG:HD3	2:EEE:32:LEU:HD22	2.00	0.43
2:DDD:58:ILE:HD13	2:DDD:58:ILE:HA	1.83	0.43
2:EEE:48:ASN:OD1	2:EEE:49:GLY:N	2.51	0.43
1:HHH:76:VAL:HG12	1:III:68:ALA:HB1	2.01	0.43
2:GGG:104:ALA:HB1	2:GGG:111:ILE:HG13	2.00	0.43
1:III:77:ASN:HD22	1:JJJ:65:SER:HB3	1.82	0.43
1:KKK:13:GLY:HA3	1:KKK:107:ARG:HH22	1.84	0.43
2:FFF:10:LEU:CD2	2:FFF:55:VAL:CA	2.97	0.43
1:LLL:14:LEU:HD11	1:LLL:32:ALA:HA	1.99	0.43
1:III:59:GLY:HA3	1:III:82:ILE:HB	2.01	0.43
2:FFF:22:ARG:NH1	2:FFF:27:LEU:O	2.52	0.43
1:HHH:2:SER:HA	1:JJJ:123:GLU:OE1	2.19	0.43
1:KKK:46:ARG:HB3	1:KKK:89:GLU:OE2	2.19	0.43
3:LLL:201:TRS:O1	3:LLL:201:TRS:O2	2.35	0.43
2:FFF:58:ILE:HD13	2:FFF:58:ILE:HA	1.94	0.43
1:III:46:ARG:HB3	1:III:89:GLU:OE2	2.19	0.43
1:LLL:67:LEU:HD12	1:LLL:67:LEU:HA	1.87	0.43
1:HHH:67:LEU:HB3	1:HHH:73:LEU:HB3	1.99	0.43
1:MMM:14:LEU:HD11	1:MMM:32:ALA:HA	2.00	0.43
1:HHH:126:GLU:HA	1:III:7:ILE:O	2.19	0.42
2:DDD:48:ASN:HD22	2:DDD:48:ASN:C	2.21	0.42
1:KKK:96:ASN:CG	1:KKK:103:ILE:CD1	2.87	0.42
1:MMM:126:GLU:N	2:AAA:56:ASN:HD21	2.18	0.42
1:LLL:15:PRO:HG2	1:LLL:30:TYR:CE2	2.55	0.42
2:CCC:18:ILE:HD13	2:CCC:61:TYR:CD2	2.54	0.42
1:HHH:48:GLY:HA2	1:HHH:89:GLU:OE1	2.20	0.42
2:EEE:94:PHE:CD1	2:EEE:94:PHE:C	2.93	0.41
1:KKK:115:LEU:N	1:KKK:115:LEU:HD12	2.35	0.41
1:MMM:109:ASP:HA	5:MMM:201:PEG:H22	2.01	0.41
2:GGG:21:LEU:O	2:GGG:25:MET:HB2	2.20	0.41
2:GGG:94:PHE:HA	2:GGG:126:LEU:HD11	2.03	0.41
1:III:98:ASP:HA	1:III:99:PRO:HD3	1.92	0.41
2:EEE:76:SER:O	2:EEE:79:PHE:HB2	2.21	0.41

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:GGG:21:LEU:O	2:GGG:25:MET:N	2.54	0.41
2:FFF:11:PRO:CG	2:FFF:13:HIS:CE1	3.03	0.41
2:AAA:69:SER:O	2:AAA:72:LEU:HB2	2.21	0.41
1:III:36:GLU:HG3	1:III:104:VAL:HG22	2.03	0.41
2:EEE:24:TYR:CD1	2:EEE:24:TYR:C	2.95	0.41
1:JJJ:9:ARG:NH1	1:JJJ:14:LEU:O	2.53	0.40
2:CCC:48:ASN:ND2	2:CCC:50:ASN:H	2.20	0.40
2:GGG:116:TYR:CD1	2:GGG:116:TYR:C	2.94	0.40
1:III:14:LEU:HD11	1:III:32:ALA:HA	2.03	0.40
1:MMM:115:LEU:HD12	1:MMM:115:LEU:N	2.35	0.40
1:HHH:8:VAL:HG23	1:JJJ:130:PHE:CZ	2.56	0.40

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	HHH	132/136 (97%)	127 (96%)	5 (4%)	0	100	100
1	III	129/136 (95%)	122 (95%)	7 (5%)	0	100	100
1	JJJ	129/136 (95%)	125 (97%)	4 (3%)	0	100	100
1	KKK	133/136 (98%)	131 (98%)	2 (2%)	0	100	100
1	LLL	134/136 (98%)	128 (96%)	6 (4%)	0	100	100
1	MMM	129/136 (95%)	116 (90%)	13 (10%)	0	100	100
2	AAA	102/147 (69%)	90 (88%)	11 (11%)	1 (1%)	15	27
2	CCC	144/147 (98%)	138 (96%)	6 (4%)	0	100	100
2	DDD	143/147 (97%)	136 (95%)	7 (5%)	0	100	100
2	EEE	144/147 (98%)	138 (96%)	6 (4%)	0	100	100
2	FFF	146/147 (99%)	142 (97%)	4 (3%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	GGG	122/147 (83%)	110 (90%)	12 (10%)	0	100	100
All	All	1587/1698 (94%)	1503 (95%)	83 (5%)	1 (0%)	51	75

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	AAA	68	PRO

### 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	HHH	92/106 (87%)	87 (95%)	5 (5%)	22	38
1	III	98/106 (92%)	84 (86%)	14 (14%)	3	4
1	JJJ	98/106 (92%)	93 (95%)	5 (5%)	24	41
1	KKK	90/106 (85%)	81 (90%)	9 (10%)	7	13
1	LLL	98/106 (92%)	94 (96%)	4 (4%)	30	50
1	MMM	99/106 (93%)	90 (91%)	9 (9%)	9	16
2	AAA	51/135 (38%)	43 (84%)	8 (16%)	2	3
2	CCC	122/135 (90%)	109 (89%)	13 (11%)	6	11
2	DDD	117/135 (87%)	110 (94%)	7 (6%)	19	33
2	EEE	111/135 (82%)	107 (96%)	4 (4%)	35	55
2	FFF	121/135 (90%)	110 (91%)	11 (9%)	9	16
2	GGG	75/135 (56%)	67 (89%)	8 (11%)	6	11
All	All	1172/1446 (81%)	1075 (92%)	97 (8%)	11	20

All (97) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	HHH	4	THR
1	HHH	41	ARG

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	HHH	44	LEU
1	HHH	45	VAL
1	HHH	110	ARG
1	III	2	SER
1	III	7	ILE
1	III	10	LEU
1	III	21	HIS
1	III	27	VAL
1	III	35	VAL
1	III	36	GLU
1	III	37	LEU
1	III	41	ARG
1	III	67	LEU
1	III	70	ARG
1	III	71	VAL
1	III	90	ILE
1	III	96	ASN
1	JJJ	33	GLU
1	JJJ	35	VAL
1	JJJ	49	VAL
1	JJJ	81	THR
1	JJJ	123	GLU
1	KKK	18	SER
1	KKK	35	VAL
1	KKK	65	SER
1	KKK	67	LEU
1	KKK	71	VAL
1	KKK	97	LEU
1	KKK	107	ARG
1	KKK	118	ARG
1	KKK	119	VAL
1	LLL	2	SER
1	LLL	4	THR
1	LLL	35	VAL
1	LLL	67	LEU
1	MMM	4	THR
1	MMM	7	ILE
1	MMM	27	VAL
1	MMM	46	ARG
1	MMM	67	LEU
1	MMM	70	ARG
1	MMM	110	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	MMM	123	GLU
1	MMM	129	SER
2	CCC	12	THR
2	CCC	40	GLN
2	CCC	48	ASN
2	CCC	58	ILE
2	CCC	72	LEU
2	CCC	76	SER
2	CCC	99	SER
2	CCC	118	LEU
2	CCC	129	LEU
2	CCC	131	GLU
2	CCC	147	LEU
2	CCC	150	GLN
2	CCC	151	ILE
2	DDD	12	THR
2	DDD	31	LYS
2	DDD	48	ASN
2	DDD	58	ILE
2	DDD	130	LYS
2	DDD	138	ASP
2	DDD	147	LEU
2	EEE	18	ILE
2	EEE	30	SER
2	EEE	42	THR
2	EEE	58	ILE
2	FFF	51	ARG
2	FFF	57	GLU
2	FFF	58	ILE
2	FFF	82	LYS
2	FFF	90	ASP
2	FFF	115	SER
2	FFF	118	LEU
2	FFF	132	SER
2	FFF	142	ASP
2	FFF	147	LEU
2	FFF	148	TYR
2	GGG	25	MET
2	GGG	27	LEU
2	GGG	36	THR
2	GGG	84	TYR
2	GGG	114	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	GGG	142	ASP
2	GGG	147	LEU
2	GGG	148	TYR
2	AAA	10	LEU
2	AAA	22	ARG
2	AAA	69	SER
2	AAA	77	ASP
2	AAA	80	LYS
2	AAA	99	SER
2	AAA	114	SER
2	AAA	148	TYR

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

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

16 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	III	201	-	3,3,3	0.10	0	2,2,2	0.27	0
4	EDO	JJJ	205	-	3,3,3	0.11	0	2,2,2	0.32	0
4	EDO	GGG	201	-	3,3,3	0.12	0	2,2,2	0.30	0
4	EDO	DDD	201	-	3,3,3	0.09	0	2,2,2	0.31	0
4	EDO	JJJ	204	-	3,3,3	0.13	0	2,2,2	0.14	0
5	PEG	JJJ	202	-	6,6,6	0.20	0	5,5,5	0.14	0
4	EDO	DDD	202	-	3,3,3	0.16	0	2,2,2	0.38	0
3	TRS	HHH	201	-	7,7,7	0.16	0	9,9,9	0.34	0
4	EDO	FFF	202	-	3,3,3	0.11	0	2,2,2	0.26	0
3	TRS	LLL	202	-	7,7,7	0.19	0	9,9,9	0.27	0
3	TRS	JJJ	201	-	7,7,7	0.24	0	9,9,9	0.32	0
5	PEG	MMM	201	-	6,6,6	0.22	0	5,5,5	0.19	0
3	TRS	LLL	201	-	7,7,7	0.18	0	9,9,9	0.28	0
5	PEG	KKK	201	-	6,6,6	0.15	0	5,5,5	0.12	0
5	PEG	FFF	201	-	6,6,6	0.17	0	5,5,5	0.08	0
4	EDO	JJJ	203	-	3,3,3	0.20	0	2,2,2	0.33	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
4	EDO	III	201	-	-	1/1/1/1	-
4	EDO	JJJ	205	-	-	1/1/1/1	-
4	EDO	GGG	201	-	-	1/1/1/1	-
4	EDO	DDD	201	-	-	1/1/1/1	-
4	EDO	JJJ	204	-	-	0/1/1/1	-
5	PEG	JJJ	202	-	-	1/4/4/4	-
4	EDO	DDD	202	-	-	1/1/1/1	-
3	TRS	HHH	201	-	-	4/9/9/9	-
4	EDO	FFF	202	-	-	1/1/1/1	-
3	TRS	LLL	202	-	-	3/9/9/9	-
3	TRS	JJJ	201	-	-	6/9/9/9	-
5	PEG	MMM	201	-	-	3/4/4/4	-
3	TRS	LLL	201	-	-	1/9/9/9	-
5	PEG	KKK	201	-	-	2/4/4/4	-
5	PEG	FFF	201	-	-	1/4/4/4	-
4	EDO	JJJ	203	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (28) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	LLL	201	TRS	N-C-C2-O2
5	MMM	201	PEG	C4-C3-O2-C2
5	KKK	201	PEG	O1-C1-C2-O2
5	KKK	201	PEG	O2-C3-C4-O4
4	III	201	EDO	O1-C1-C2-O2
4	DDD	202	EDO	O1-C1-C2-O2
4	FFF	202	EDO	O1-C1-C2-O2
3	HHH	201	TRS	C3-C-C1-O1
3	JJJ	201	TRS	C2-C-C1-O1
4	GGG	201	EDO	O1-C1-C2-O2
4	DDD	201	EDO	O1-C1-C2-O2
5	JJJ	202	PEG	C1-C2-O2-C3
5	MMM	201	PEG	O1-C1-C2-O2
3	HHH	201	TRS	N-C-C1-O1
3	JJJ	201	TRS	N-C-C1-O1
3	JJJ	201	TRS	N-C-C2-O2
3	LLL	202	TRS	C1-C-C3-O3
4	JJJ	203	EDO	O1-C1-C2-O2
3	HHH	201	TRS	C2-C-C1-O1
3	JJJ	201	TRS	C3-C-C1-O1
3	JJJ	201	TRS	C1-C-C2-O2
3	JJJ	201	TRS	C3-C-C2-O2
3	LLL	202	TRS	C2-C-C3-O3
5	FFF	201	PEG	C4-C3-O2-C2
5	MMM	201	PEG	O2-C3-C4-O4
4	JJJ	205	EDO	O1-C1-C2-O2
3	HHH	201	TRS	N-C-C2-O2
3	LLL	202	TRS	N-C-C3-O3

There are no ring outliers.

4 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	HHH	201	TRS	5	0
3	LLL	202	TRS	4	0
5	MMM	201	PEG	2	0
3	LLL	201	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

### 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	HHH	134/136 (98%)	0.06	0 <span>100</span> <span>100</span>	39, 52, 70, 118	0
1	III	131/136 (96%)	0.04	0 <span>100</span> <span>100</span>	37, 52, 90, 115	0
1	JJJ	131/136 (96%)	0.03	0 <span>100</span> <span>100</span>	40, 58, 80, 113	0
1	KKK	135/136 (99%)	0.05	1 (0%) <span>87</span> <span>91</span>	41, 68, 94, 106	0
1	LLL	136/136 (100%)	-0.11	0 <span>100</span> <span>100</span>	32, 51, 77, 123	0
1	MMM	131/136 (96%)	-0.09	0 <span>100</span> <span>100</span>	34, 56, 86, 110	0
2	AAA	118/147 (80%)	0.19	6 (5%) <span>28</span> <span>34</span>	58, 94, 136, 160	1 (0%)
2	CCC	146/147 (99%)	-0.17	0 <span>100</span> <span>100</span>	35, 51, 81, 104	0
2	DDD	145/147 (98%)	-0.14	1 (0%) <span>87</span> <span>91</span>	40, 61, 91, 105	0
2	EEE	146/147 (99%)	-0.11	1 (0%) <span>87</span> <span>91</span>	44, 63, 98, 116	0
2	FFF	146/147 (99%)	0.00	1 (0%) <span>87</span> <span>91</span>	38, 54, 90, 116	0
2	GGG	128/147 (87%)	-0.13	0 <span>100</span> <span>100</span>	50, 81, 124, 140	0
All	All	1627/1698 (95%)	-0.04	10 (0%) <span>89</span> <span>92</span>	32, 59, 103, 160	1 (0%)

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	AAA	139	ILE	4.5
2	AAA	75	ILE	3.7
2	FFF	12	THR	3.2
2	AAA	10	LEU	3.1
2	AAA	30	SER	3.0
2	DDD	148	TYR	2.9
2	AAA	13	HIS	2.6
2	AAA	123	ILE	2.5
1	KKK	71	VAL	2.1
2	EEE	116	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
4	EDO	DDD	202	4/4	0.57	0.25	93,96,96,96	0
4	EDO	JJJ	203	4/4	0.69	0.25	75,85,88,89	0
4	EDO	GGG	201	4/4	0.76	0.18	94,95,96,98	0
5	PEG	MMM	201	7/7	0.80	0.28	98,108,115,119	0
5	PEG	KKK	201	7/7	0.82	0.19	76,87,104,105	0
4	EDO	JJJ	205	4/4	0.82	0.11	93,98,99,100	0
5	PEG	FFF	201	7/7	0.87	0.26	86,97,101,103	0
5	PEG	JJJ	202	7/7	0.89	0.10	98,104,108,111	0
4	EDO	JJJ	204	4/4	0.89	0.12	95,102,104,106	0
4	EDO	FFF	202	4/4	0.92	0.13	71,76,79,82	0
4	EDO	III	201	4/4	0.92	0.24	76,80,82,82	0
3	TRS	HHH	201	8/8	0.95	0.25	63,65,68,73	0
4	EDO	DDD	201	4/4	0.96	0.22	65,68,71,72	0
3	TRS	LLL	201	8/8	0.96	0.25	70,76,81,85	0
3	TRS	LLL	202	8/8	0.96	0.25	55,59,61,62	0
3	TRS	JJJ	201	8/8	0.96	0.28	62,70,78,78	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.