



# Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 03:46 pm BST

PDB ID : 2F8J  
Title : Crystal structure of Histidinol-phosphate aminotransferase (EC 2.6.1.9) (Imidazole acetol-phosphate transferase) (tm1040) from *Thermotoga maritima* at 2.40 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2005-12-02  
Resolution : 2.40 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 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

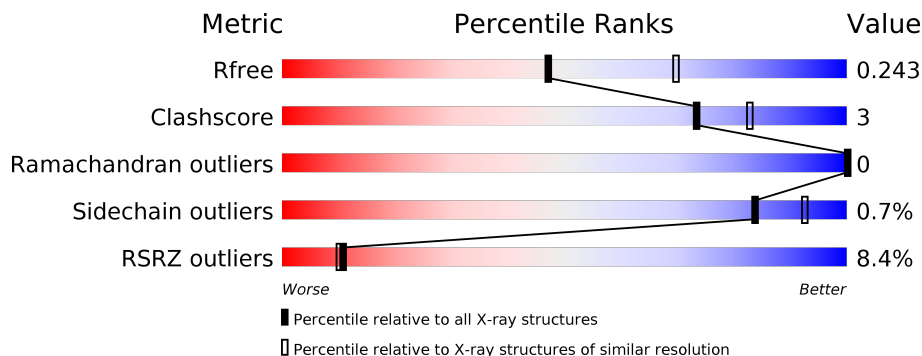
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	347	 3% 86% 10% ..
1	B	347	 6% 88% 7% .
1	C	347	 6% 89% 7% ..
1	D	347	 16% 91% 5% .

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 11003 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 Histidinol-phosphate aminotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	335	Total	C	N	O	S	0	1	0
			2685	1741	451	484	9			
1	B	333	Total	C	N	O	S	0	0	0
			2636	1709	443	475	9			
1	C	335	Total	C	N	O	S	0	3	0
			2695	1750	449	485	11			
1	D	333	Total	C	N	O	S	0	0	0
			2532	1641	411	472	8			

There are 48 discrepancies between the modelled and reference sequences:

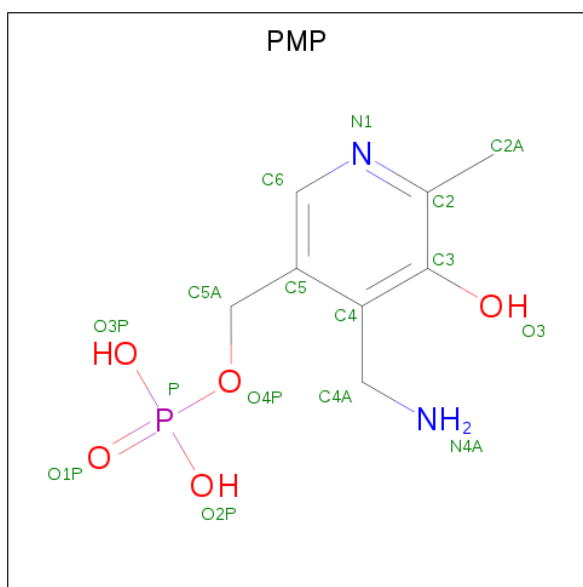
Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	LEADER SEQUENCE	UNP Q9X0D0
A	-10	GLY	-	LEADER SEQUENCE	UNP Q9X0D0
A	-9	SER	-	LEADER SEQUENCE	UNP Q9X0D0
A	-8	ASP	-	LEADER SEQUENCE	UNP Q9X0D0
A	-7	LYS	-	LEADER SEQUENCE	UNP Q9X0D0
A	-6	ILE	-	LEADER SEQUENCE	UNP Q9X0D0
A	-5	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
A	-4	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
A	-3	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
A	-2	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
A	-1	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
A	0	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
B	-11	MET	-	LEADER SEQUENCE	UNP Q9X0D0
B	-10	GLY	-	LEADER SEQUENCE	UNP Q9X0D0
B	-9	SER	-	LEADER SEQUENCE	UNP Q9X0D0
B	-8	ASP	-	LEADER SEQUENCE	UNP Q9X0D0
B	-7	LYS	-	LEADER SEQUENCE	UNP Q9X0D0
B	-6	ILE	-	LEADER SEQUENCE	UNP Q9X0D0
B	-5	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
B	-4	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
B	-3	HIS	-	LEADER SEQUENCE	UNP Q9X0D0

*Continued on next page...*

*Continued from previous page...*

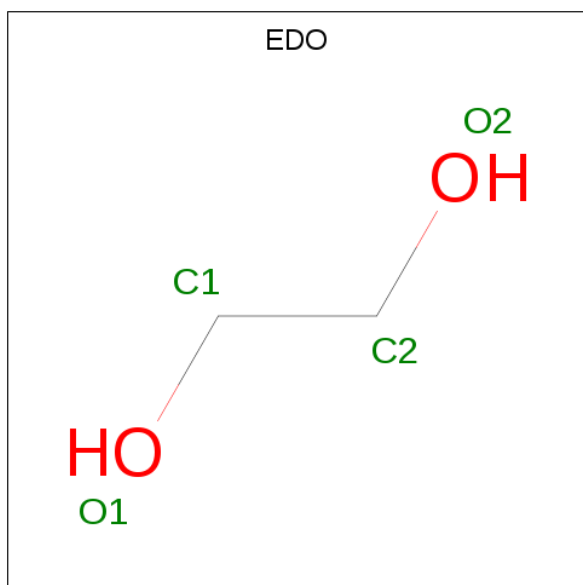
Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
B	-1	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
B	0	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
C	-11	MET	-	LEADER SEQUENCE	UNP Q9X0D0
C	-10	GLY	-	LEADER SEQUENCE	UNP Q9X0D0
C	-9	SER	-	LEADER SEQUENCE	UNP Q9X0D0
C	-8	ASP	-	LEADER SEQUENCE	UNP Q9X0D0
C	-7	LYS	-	LEADER SEQUENCE	UNP Q9X0D0
C	-6	ILE	-	LEADER SEQUENCE	UNP Q9X0D0
C	-5	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
C	-4	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
C	-3	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
C	-2	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
C	-1	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
C	0	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
D	-11	MET	-	LEADER SEQUENCE	UNP Q9X0D0
D	-10	GLY	-	LEADER SEQUENCE	UNP Q9X0D0
D	-9	SER	-	LEADER SEQUENCE	UNP Q9X0D0
D	-8	ASP	-	LEADER SEQUENCE	UNP Q9X0D0
D	-7	LYS	-	LEADER SEQUENCE	UNP Q9X0D0
D	-6	ILE	-	LEADER SEQUENCE	UNP Q9X0D0
D	-5	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
D	-4	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
D	-3	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
D	-2	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
D	-1	HIS	-	LEADER SEQUENCE	UNP Q9X0D0
D	0	HIS	-	LEADER SEQUENCE	UNP Q9X0D0

- Molecule 2 is 4'-DEOXY-4'-AMINOPYRIDOXAL-5'-PHOSPHATE (three-letter code: PMP) (formula: C<sub>8</sub>H<sub>13</sub>N<sub>2</sub>O<sub>5</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	Total 16	8	2	5	1	0	0
2	B	1	Total 16	8	2	5	1	0	0
2	C	1	Total 16	8	2	5	1	0	0
2	D	1	Total 16	8	2	5	1	0	0

- 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	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0
3	D	1	Total C O 4 2 2	0	0
3	D	1	Total C O 4 2 2	0	0

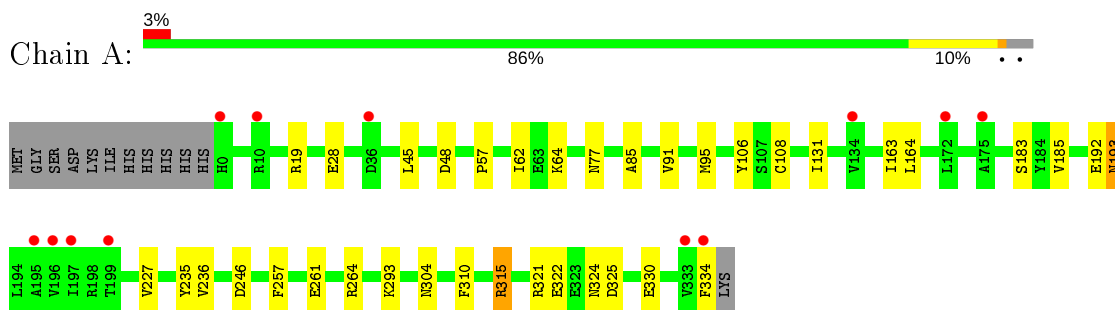
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	100	Total O 100 100	0	0
4	B	110	Total O 110 110	0	0
4	C	98	Total O 98 98	0	0
4	D	31	Total O 31 31	0	0

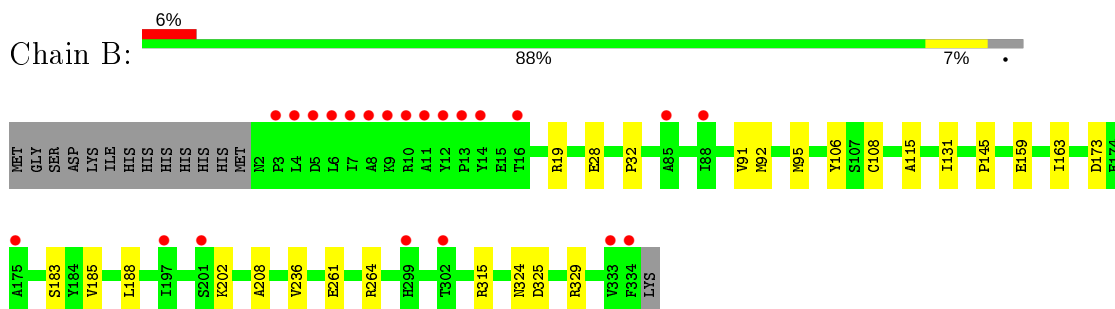
### 3 Residue-property plots [i](#)

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.

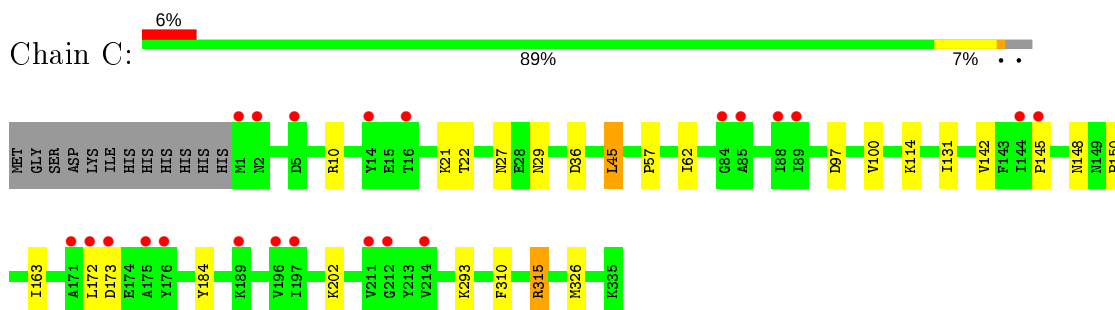
- Molecule 1: Histidinol-phosphate aminotransferase



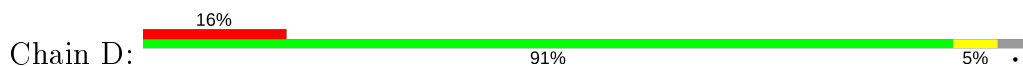
- Molecule 1: Histidinol-phosphate aminotransferase

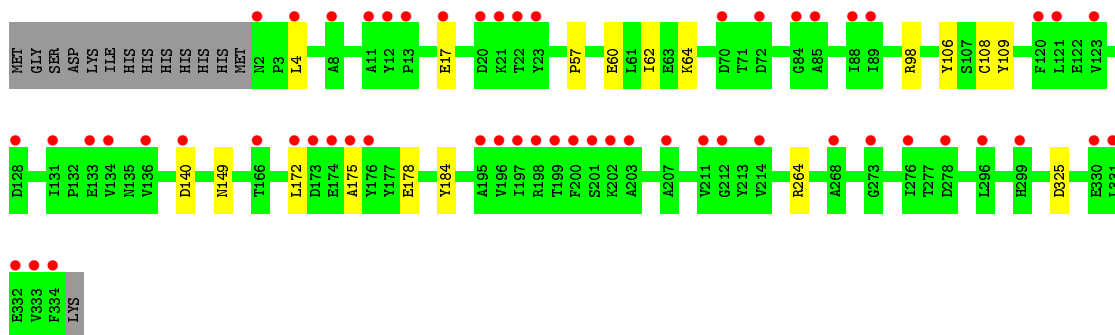


- Molecule 1: Histidinol-phosphate aminotransferase



- Molecule 1: Histidinol-phosphate aminotransferase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	145.16Å 187.45Å 54.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.60 – 2.40 29.79 – 2.39	Depositor EDS
% Data completeness (in resolution range)	91.5 (29.60-2.40) 91.5 (29.79-2.39)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.60 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.182 , 0.238 0.189 , 0.243	Depositor DCC
$R_{free}$ test set	2745 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.0	Xtrriage
Anisotropy	0.240	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 57.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	11003	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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: PMP, 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	0.55	0/2745	0.67	2/3715 (0.1%)
1	B	0.56	0/2695	0.65	0/3645
1	C	0.54	0/2761	0.68	3/3730 (0.1%)
1	D	0.44	0/2591	0.57	0/3515
All	All	0.53	0/10792	0.64	5/14605 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	C	315	ARG	NE-CZ-NH1	6.43	123.52	120.30
1	C	45	LEU	CA-CB-CG	6.13	129.41	115.30
1	A	315	ARG	NE-CZ-NH1	6.11	123.36	120.30
1	C	315	ARG	NE-CZ-NH2	-5.55	117.52	120.30
1	A	193	ASN	N-CA-CB	-5.03	101.55	110.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	192	GLU	Peptide

## 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	A	2685	0	2601	24	0
1	B	2636	0	2533	23	0
1	C	2695	0	2625	16	0
1	D	2532	0	2307	10	0
2	A	16	0	11	1	0
2	B	16	0	10	3	0
2	C	16	0	10	1	0
2	D	16	0	10	2	0
3	A	12	0	18	0	0
3	B	20	0	30	0	0
3	C	12	0	18	0	0
3	D	8	0	12	2	0
4	A	100	0	0	1	0
4	B	110	0	0	2	0
4	C	98	0	0	1	0
4	D	31	0	0	0	0
All	All	11003	0	10185	69	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.

All (69) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:131:ILE:HG21	1:C:163:ILE:HD11	1.69	0.74
1:C:10:ARG:NH1	1:C:114:LYS:O	2.25	0.69
1:B:131:ILE:HG21	1:B:163:ILE:HD11	1.82	0.61
1:A:261:GLU:OE2	1:A:321:ARG:HD3	2.01	0.60
1:D:149:ASN:ND2	2:D:500:PMP:O3	2.35	0.59
1:A:45:LEU:HD11	1:A:236:VAL:HG22	1.85	0.58
1:B:183:SER:OG	1:B:185:VAL:HG23	2.03	0.58
1:A:261:GLU:HG3	1:A:324:ASN:ND2	2.19	0.57
1:C:315:ARG:HD2	4:C:520:HOH:O	2.04	0.56
1:C:172:LEU:CD2	1:C:184:TYR:HB2	2.36	0.56
1:C:202:LYS:NZ	2:C:500:PMP:H4A2	2.22	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:178:GLU:H	3:D:502:EDO:H22	1.72	0.54
1:A:57:PRO:HG3	1:A:62:ILE:HD11	1.88	0.54
1:D:172:LEU:CD2	1:D:184:TYR:HB2	2.37	0.54
1:B:159:GLU:O	1:B:163:ILE:HD12	2.08	0.54
1:A:261:GLU:CG	1:A:324:ASN:ND2	2.70	0.53
1:A:236:VAL:HG11	1:B:236:VAL:HG11	1.89	0.53
1:A:235:TYR:HB2	1:B:208:ALA:HB1	1.90	0.52
1:C:100:VAL:HB	1:C:142:VAL:HG22	1.92	0.52
1:C:131:ILE:HG21	1:C:163:ILE:CD1	2.37	0.51
1:D:57:PRO:HG3	1:D:62:ILE:HD11	1.91	0.51
1:A:315:ARG:HD2	4:A:506:HOH:O	2.09	0.51
1:A:164:LEU:O	1:A:193:ASN:ND2	2.43	0.50
1:B:261:GLU:HG2	1:B:324:ASN:ND2	2.26	0.50
1:B:261:GLU:CG	1:B:324:ASN:ND2	2.76	0.49
1:D:106:TYR:CE2	1:D:108:CYS:HB2	2.46	0.49
1:D:175:ALA:O	3:D:502:EDO:O2	2.25	0.49
1:A:293:LYS:NZ	1:A:310:PHE:O	2.38	0.49
1:B:202:LYS:NZ	2:B:500:PMP:C4A	2.76	0.49
1:C:293:LYS:NZ	1:C:310:PHE:O	2.43	0.48
1:C:57:PRO:HG3	1:C:62:ILE:HD11	1.96	0.48
1:A:106:TYR:CE2	1:A:108:CYS:HB2	2.50	0.47
1:A:257:PHE:O	1:A:261:GLU:HB2	2.14	0.47
1:A:57:PRO:CG	1:A:62:ILE:HD11	2.45	0.47
1:A:183:SER:OG	1:A:185:VAL:HG23	2.15	0.46
1:B:19:ARG:NH2	1:B:28:GLU:OE2	2.48	0.46
1:B:202:LYS:HZ1	2:B:500:PMP:C4A	2.30	0.45
1:C:22:THR:HG21	1:C:326:MET:HE1	1.96	0.45
1:B:92:MET:HG2	1:B:95:MET:HE3	1.99	0.45
1:B:261:GLU:CG	1:B:324:ASN:HD22	2.30	0.45
1:A:85:ALA:HB3	2:A:500:PMP:H5A2	1.99	0.45
1:A:91:VAL:O	1:A:95:MET:HG3	2.17	0.44
1:A:330:GLU:O	1:A:334:PHE:HB2	2.17	0.44
1:D:106:TYR:HB3	1:D:109:TYR:CE2	2.52	0.44
1:C:145:PRO:HB2	1:C:148:ASN:HA	1.98	0.44
1:A:48:ASP:HA	1:B:32:PRO:HB3	1.99	0.44
1:B:315:ARG:HD2	4:B:514:HOH:O	2.17	0.43
1:C:29:ASN:HB2	1:C:202:LYS:O	2.19	0.43
1:A:131:ILE:HG21	1:A:163:ILE:HD11	2.01	0.43
1:A:77:ASN:ND2	1:C:97:ASP:OD2	2.51	0.43
1:A:64:LYS:NZ	1:A:246:ASP:OD1	2.50	0.43
1:B:329:ARG:NH2	4:B:520:HOH:O	2.51	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:98:ARG:HB3	1:D:140:ASP:OD1	2.19	0.42
1:B:91:VAL:O	1:B:95:MET:HG3	2.18	0.42
1:C:145:PRO:HA	1:C:173:ASP:HB3	2.02	0.42
1:C:27:ASN:O	1:C:202:LYS:HD2	2.19	0.42
2:D:500:PMP:O3	2:D:500:PMP:N4A	2.51	0.42
1:C:150:PRO:HB3	1:C:310:PHE:CZ	2.53	0.42
1:B:106:TYR:CE2	1:B:108:CYS:HB2	2.55	0.42
1:B:185:VAL:HG13	1:B:188:LEU:HD12	2.02	0.42
1:A:227:VAL:HA	1:B:115:ALA:HB1	2.02	0.42
1:B:202:LYS:HZ1	2:B:500:PMP:H4A2	1.85	0.42
1:D:60:GLU:HG2	1:D:64:LYS:HE2	2.01	0.41
1:A:19[A]:ARG:NH2	1:A:28:GLU:OE2	2.53	0.41
1:B:264:ARG:NH2	1:B:325:ASP:OD1	2.54	0.41
1:B:131:ILE:HD13	1:B:163:ILE:HD11	2.02	0.41
1:A:264:ARG:NH2	1:A:325:ASP:OD1	2.55	0.40
1:D:264:ARG:NH2	1:D:325:ASP:OD1	2.54	0.40
1:B:145:PRO:HA	1:B:173:ASP:HB3	2.04	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	A	334/347 (96%)	327 (98%)	7 (2%)	0	100	100
1	B	331/347 (95%)	321 (97%)	10 (3%)	0	100	100
1	C	336/347 (97%)	328 (98%)	8 (2%)	0	100	100
1	D	331/347 (95%)	322 (97%)	9 (3%)	0	100	100
All	All	1332/1388 (96%)	1298 (97%)	34 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	272/312 (87%)	270 (99%)	2 (1%)	84	92
1	B	264/312 (85%)	264 (100%)	0	100	100
1	C	273/312 (88%)	270 (99%)	3 (1%)	73	87
1	D	238/312 (76%)	236 (99%)	2 (1%)	81	91
All	All	1047/1248 (84%)	1040 (99%)	7 (1%)	84	92

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	304	ASN
1	A	322	GLU
1	C	21	LYS
1	C	36	ASP
1	C	45	LEU
1	D	4	LEU
1	D	17	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	A	77	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

17 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	A	501	-	3,3,3	0.60	0	2,2,2	0.18	0
3	EDO	C	501	-	3,3,3	0.64	0	2,2,2	0.09	0
3	EDO	B	503	-	3,3,3	0.55	0	2,2,2	0.18	0
2	PMP	C	500	-	16,16,16	1.50	2 (12%)	21,23,23	1.85	7 (33%)
2	PMP	D	500	-	16,16,16	1.45	2 (12%)	21,23,23	1.79	6 (28%)
3	EDO	B	504	-	3,3,3	0.51	0	2,2,2	0.13	0
3	EDO	B	505	-	3,3,3	0.43	0	2,2,2	0.35	0
2	PMP	A	500	-	16,16,16	1.43	3 (18%)	21,23,23	1.97	6 (28%)
2	PMP	B	500	-	16,16,16	1.37	2 (12%)	21,23,23	1.83	5 (23%)
3	EDO	B	501	-	3,3,3	0.60	0	2,2,2	0.35	0
3	EDO	D	501	-	3,3,3	0.65	0	2,2,2	0.03	0
3	EDO	C	503	-	3,3,3	0.42	0	2,2,2	0.23	0
3	EDO	A	503	-	3,3,3	0.45	0	2,2,2	0.36	0
3	EDO	A	502	-	3,3,3	0.44	0	2,2,2	0.53	0
3	EDO	D	502	-	3,3,3	0.62	0	2,2,2	0.18	0
3	EDO	B	502	-	3,3,3	0.58	0	2,2,2	0.18	0
3	EDO	C	502	-	3,3,3	0.48	0	2,2,2	0.27	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	A	501	-	-	1/1/1/1	-
3	EDO	C	501	-	-	1/1/1/1	-

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	B	503	-	-	1/1/1/1	-
2	PMP	C	500	-	-	5/8/8/8	0/1/1/1
2	PMP	D	500	-	-	4/8/8/8	0/1/1/1
3	EDO	B	504	-	-	0/1/1/1	-
3	EDO	B	505	-	-	1/1/1/1	-
2	PMP	A	500	-	-	4/8/8/8	0/1/1/1
2	PMP	B	500	-	-	3/8/8/8	0/1/1/1
3	EDO	B	501	-	-	0/1/1/1	-
3	EDO	D	501	-	-	0/1/1/1	-
3	EDO	C	503	-	-	0/1/1/1	-
3	EDO	A	503	-	-	1/1/1/1	-
3	EDO	A	502	-	-	1/1/1/1	-
3	EDO	D	502	-	-	1/1/1/1	-
3	EDO	B	502	-	-	1/1/1/1	-
3	EDO	C	502	-	-	1/1/1/1	-

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	500	PMP	P-O1P	3.21	1.60	1.50
2	D	500	PMP	P-O1P	3.20	1.60	1.50
2	A	500	PMP	P-O1P	3.14	1.60	1.50
2	B	500	PMP	P-O1P	3.04	1.60	1.50
2	A	500	PMP	C2A-C2	2.76	1.55	1.50
2	D	500	PMP	C2A-C2	2.47	1.54	1.50
2	C	500	PMP	C2A-C2	2.21	1.54	1.50
2	A	500	PMP	O4P-C5A	-2.15	1.37	1.45
2	B	500	PMP	C2A-C2	2.12	1.54	1.50

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	500	PMP	O4P-C5A-C5	4.48	117.89	109.35
2	B	500	PMP	O4P-C5A-C5	4.34	117.62	109.35
2	A	500	PMP	C6-C5-C4	4.10	121.02	118.12
2	D	500	PMP	O4P-C5A-C5	3.84	116.67	109.35
2	C	500	PMP	O4P-C5A-C5	3.72	116.44	109.35
2	C	500	PMP	O3P-P-O4P	3.48	115.98	106.73
2	D	500	PMP	O3P-P-O4P	3.47	115.97	106.73
2	B	500	PMP	O2P-P-O1P	-3.19	98.21	110.68
2	A	500	PMP	O3P-P-O4P	3.11	115.00	106.73

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	500	PMP	C6-C5-C4	3.03	120.26	118.12
2	C	500	PMP	O2P-P-O1P	-2.82	99.65	110.68
2	B	500	PMP	O3P-P-O4P	2.76	114.07	106.73
2	C	500	PMP	C3-C4-C5	-2.65	116.18	118.72
2	C	500	PMP	C6-C5-C4	2.53	119.91	118.12
2	D	500	PMP	O2P-P-O1P	-2.49	100.92	110.68
2	A	500	PMP	C5-C6-N1	-2.47	119.71	123.82
2	B	500	PMP	C2A-C2-N1	2.41	122.37	117.67
2	B	500	PMP	C2A-C2-C3	-2.22	118.14	120.89
2	A	500	PMP	O2P-P-O4P	-2.20	100.87	106.73
2	D	500	PMP	C5A-C5-C6	-2.20	115.75	119.37
2	D	500	PMP	C3-C4-C5	-2.10	116.70	118.72
2	A	500	PMP	C6-N1-C2	2.09	123.03	119.17
2	C	500	PMP	O2P-P-O4P	-2.05	101.27	106.73
2	C	500	PMP	C4A-C4-C3	2.01	123.54	120.34

There are no chirality outliers.

All (25) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	500	PMP	C5A-O4P-P-O2P
2	C	500	PMP	C5A-O4P-P-O3P
2	D	500	PMP	C5A-O4P-P-O2P
2	A	500	PMP	C5A-O4P-P-O2P
2	A	500	PMP	C5A-O4P-P-O3P
2	B	500	PMP	C5A-O4P-P-O1P
2	B	500	PMP	C5A-O4P-P-O2P
2	B	500	PMP	C5A-O4P-P-O3P
2	C	500	PMP	C3-C4-C4A-N4A
2	D	500	PMP	C3-C4-C4A-N4A
3	B	503	EDO	O1-C1-C2-O2
2	C	500	PMP	C5-C4-C4A-N4A
3	C	501	EDO	O1-C1-C2-O2
2	C	500	PMP	C5A-O4P-P-O1P
2	A	500	PMP	C5A-O4P-P-O1P
2	D	500	PMP	C5-C4-C4A-N4A
3	B	502	EDO	O1-C1-C2-O2
3	A	503	EDO	O1-C1-C2-O2
3	C	502	EDO	O1-C1-C2-O2
3	A	501	EDO	O1-C1-C2-O2
3	D	502	EDO	O1-C1-C2-O2
2	D	500	PMP	C5A-O4P-P-O3P

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	B	505	EDO	O1-C1-C2-O2
3	A	502	EDO	O1-C1-C2-O2
2	A	500	PMP	C5-C4-C4A-N4A

There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	500	PMP	1	0
2	D	500	PMP	2	0
2	A	500	PMP	1	0
2	B	500	PMP	3	0
3	D	502	EDO	2	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	A	335/347 (96%)	-0.06	12 (3%) 42 42	48, 57, 70, 93	0
1	B	333/347 (95%)	0.06	22 (6%) 18 17	48, 57, 72, 84	0
1	C	335/347 (96%)	0.03	22 (6%) 18 17	49, 57, 71, 85	0
1	D	333/347 (95%)	0.63	56 (16%) 1 1	48, 57, 70, 87	0
All	All	1336/1388 (96%)	0.17	112 (8%) 11 10	48, 57, 71, 93	0

All (112) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	273	GLY	5.8
1	D	334	PHE	5.4
1	C	197	ILE	4.7
1	D	333	VAL	4.7
1	D	20	ASP	4.6
1	D	136	VAL	4.6
1	D	175	ALA	4.4
1	D	12	TYR	4.3
1	D	199	THR	4.3
1	B	8	ALA	4.2
1	D	201	SER	4.2
1	B	12	TYR	4.2
1	A	334	PHE	4.2
1	B	9	LYS	4.0
1	B	5	ASP	4.0
1	D	197	ILE	3.9
1	D	72	ASP	3.8
1	B	13	PRO	3.8
1	B	4	LEU	3.8
1	C	175	ALA	3.7
1	A	175	ALA	3.7

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	334	PHE	3.7
1	B	14	TYR	3.7
1	B	6	LEU	3.6
1	B	333	VAL	3.6
1	C	171	ALA	3.6
1	A	0	HIS	3.6
1	C	1	MET	3.6
1	A	197	ILE	3.6
1	C	172	LEU	3.6
1	B	16	THR	3.5
1	D	84	GLY	3.4
1	D	166	THR	3.4
1	D	296	LEU	3.3
1	D	331	LEU	3.2
1	D	17	GLU	3.2
1	D	134	VAL	3.2
1	D	202	LYS	3.2
1	D	11	ALA	3.1
1	D	299	HIS	3.1
1	D	268	ALA	3.1
1	C	145	PRO	3.1
1	B	10	ARG	3.1
1	A	172	LEU	3.1
1	D	4	LEU	3.1
1	B	7	ILE	3.0
1	C	14	TYR	3.0
1	C	88	ILE	3.0
1	B	11	ALA	2.9
1	D	173	ASP	2.9
1	D	13	PRO	2.9
1	B	175	ALA	2.8
1	D	2	ASN	2.8
1	C	214	VAL	2.8
1	D	176	TYR	2.8
1	D	276	ILE	2.8
1	A	134	VAL	2.8
1	C	173	ASP	2.8
1	C	84	GLY	2.7
1	D	121	LEU	2.7
1	B	302	THR	2.7
1	D	203	ALA	2.6
1	D	198	ARG	2.6

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	D	22	THR	2.6
1	D	195	ALA	2.6
1	D	128	ASP	2.6
1	A	196	VAL	2.6
1	A	333	VAL	2.6
1	C	85	ALA	2.6
1	D	207	ALA	2.6
1	D	70	ASP	2.6
1	D	21	LYS	2.5
1	D	172	LEU	2.5
1	B	201	SER	2.5
1	D	133	GLU	2.5
1	A	199	THR	2.5
1	D	200	PHE	2.5
1	C	89	ILE	2.4
1	D	85	ALA	2.4
1	C	5	ASP	2.4
1	C	16	THR	2.4
1	D	88	ILE	2.3
1	C	211	VAL	2.3
1	D	8	ALA	2.3
1	D	212	GLY	2.3
1	D	89	ILE	2.3
1	D	131	ILE	2.3
1	C	176	TYR	2.3
1	D	123	VAL	2.2
1	D	140	ASP	2.2
1	D	174	GLU	2.2
1	D	330	GLU	2.2
1	B	3	PRO	2.2
1	A	10	ARG	2.2
1	D	196	VAL	2.2
1	A	195	ALA	2.2
1	B	85	ALA	2.2
1	C	144	ILE	2.2
1	D	214	VAL	2.2
1	A	36	ASP	2.2
1	D	278	ASP	2.2
1	D	23	TYR	2.1
1	B	88	ILE	2.1
1	B	299	HIS	2.1
1	C	196	VAL	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	211	VAL	2.1
1	D	120	PHE	2.0
1	D	332	GLU	2.0
1	C	2	ASN	2.0
1	B	197	ILE	2.0
1	C	212	GLY	2.0
1	C	189	LYS	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 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(Å <sup>2</sup> )	Q<0.9
3	EDO	A	501	4/4	0.74	0.26	78,78,79,83	0
3	EDO	A	502	4/4	0.85	0.45	65,71,74,77	0
3	EDO	B	502	4/4	0.86	0.30	76,78,80,81	0
3	EDO	B	501	4/4	0.88	0.16	48,60,65,72	0
3	EDO	A	503	4/4	0.88	0.21	72,76,76,80	0
3	EDO	C	501	4/4	0.89	0.23	63,66,67,68	0
3	EDO	D	502	4/4	0.89	0.41	58,67,68,74	0
2	PMP	D	500	16/16	0.89	0.21	69,90,92,92	0
3	EDO	B	504	4/4	0.90	0.36	68,72,73,73	0
2	PMP	A	500	16/16	0.93	0.18	46,71,79,79	0
3	EDO	B	503	4/4	0.93	0.15	75,76,78,79	0
2	PMP	C	500	16/16	0.93	0.17	55,69,73,74	0
3	EDO	D	501	4/4	0.94	0.15	61,68,69,70	0
3	EDO	B	505	4/4	0.95	0.27	47,49,52,53	0
2	PMP	B	500	16/16	0.95	0.16	40,68,76,86	0
3	EDO	C	502	4/4	0.95	0.08	73,75,75,75	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	C	503	4/4	0.96	0.26	47,50,52,54	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.