



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2023 – 02:11 PM EDT

PDB ID : 6CBK  
Title : X-ray structure of NeoB from *Streptomyces fradiae* in complex with PMP  
Authors : Thoden, J.B.; Dow, G.T.; Holden, H.M.  
Deposited on : 2018-02-03  
Resolution : 1.75 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

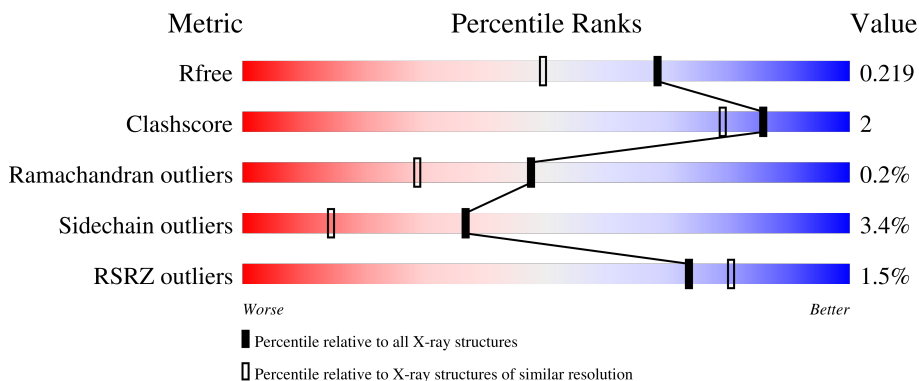
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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

Mol	Chain	Length	Quality of chain
1	A	424	 90% 5% . .
1	B	424	 91% 8% .
1	C	424	 90% 5% . .
1	D	424	 91% 7% .

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 13702 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 Neamine transaminase NeoN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	409	Total 3120	C 1972	N 566	O 574	S 8	0	4	0
1	B	419	Total 3206	C 2021	N 588	O 589	S 8	0	3	0
1	C	410	Total 3103	C 1958	N 561	O 576	S 8	0	1	0
1	D	422	Total 3233	C 2036	N 589	O 600	S 8	0	5	0

There are 32 discrepancies between the modelled and reference sequences:

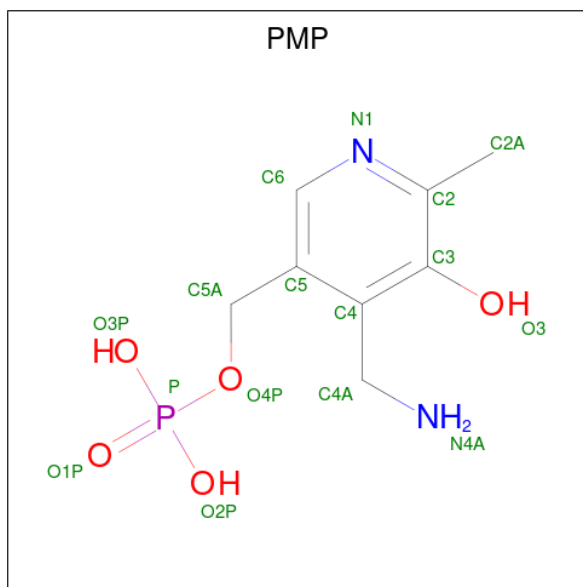
Chain	Residue	Modelled	Actual	Comment	Reference
A	417	LEU	-	expression tag	UNP Q53U08
A	418	GLU	-	expression tag	UNP Q53U08
A	419	HIS	-	expression tag	UNP Q53U08
A	420	HIS	-	expression tag	UNP Q53U08
A	421	HIS	-	expression tag	UNP Q53U08
A	422	HIS	-	expression tag	UNP Q53U08
A	423	HIS	-	expression tag	UNP Q53U08
A	424	HIS	-	expression tag	UNP Q53U08
B	417	LEU	-	expression tag	UNP Q53U08
B	418	GLU	-	expression tag	UNP Q53U08
B	419	HIS	-	expression tag	UNP Q53U08
B	420	HIS	-	expression tag	UNP Q53U08
B	421	HIS	-	expression tag	UNP Q53U08
B	422	HIS	-	expression tag	UNP Q53U08
B	423	HIS	-	expression tag	UNP Q53U08
B	424	HIS	-	expression tag	UNP Q53U08
C	417	LEU	-	expression tag	UNP Q53U08
C	418	GLU	-	expression tag	UNP Q53U08
C	419	HIS	-	expression tag	UNP Q53U08
C	420	HIS	-	expression tag	UNP Q53U08
C	421	HIS	-	expression tag	UNP Q53U08

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	422	HIS	-	expression tag	UNP Q53U08
C	423	HIS	-	expression tag	UNP Q53U08
C	424	HIS	-	expression tag	UNP Q53U08
D	417	LEU	-	expression tag	UNP Q53U08
D	418	GLU	-	expression tag	UNP Q53U08
D	419	HIS	-	expression tag	UNP Q53U08
D	420	HIS	-	expression tag	UNP Q53U08
D	421	HIS	-	expression tag	UNP Q53U08
D	422	HIS	-	expression tag	UNP Q53U08
D	423	HIS	-	expression tag	UNP Q53U08
D	424	HIS	-	expression tag	UNP Q53U08

- 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	C	N	O	P	0	0
			16	8	2	5	1		
2	B	1	Total	C	N	O	P	0	0
			16	8	2	5	1		
2	C	1	Total	C	N	O	P	0	0
			16	8	2	5	1		
2	D	1	Total	C	N	O	P	0	0
			16	8	2	5	1		

- 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

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	1	Total Na 1 1	0	0

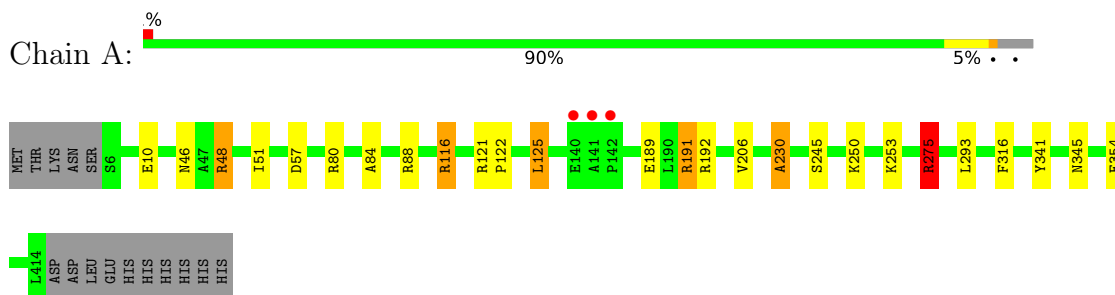
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	252	Total O 252 252	0	0
5	B	240	Total O 240 240	0	0
5	C	203	Total O 203 203	0	0
5	D	276	Total O 276 276	0	0

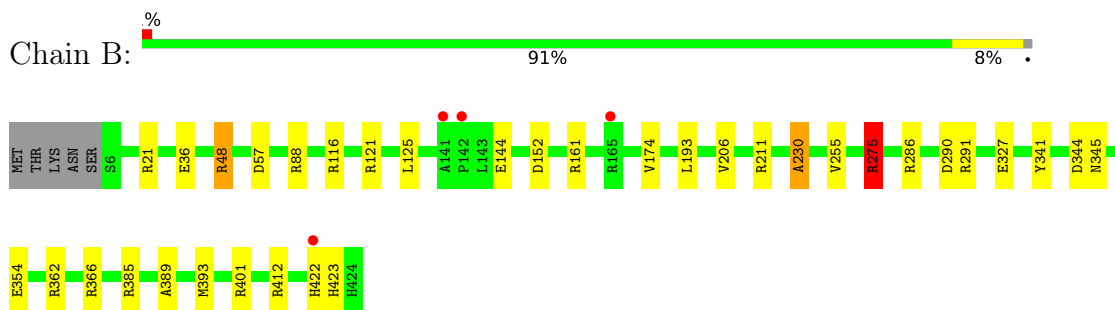
### 3 Residue-property plots [i](#)

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.

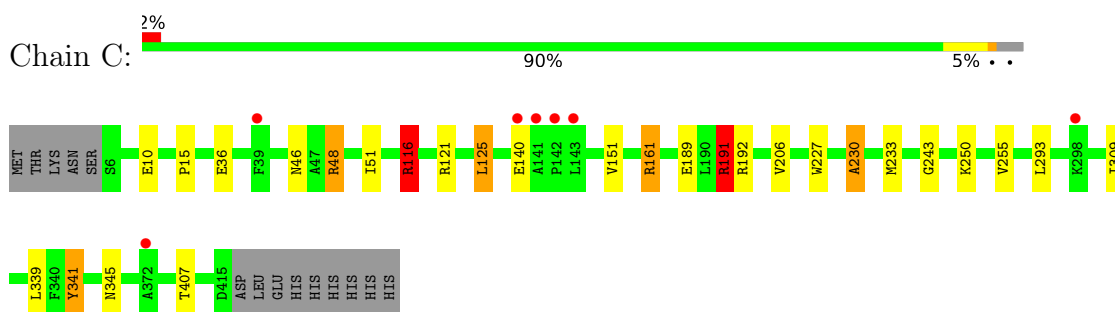
- Molecule 1: Neamine transaminase NeoN



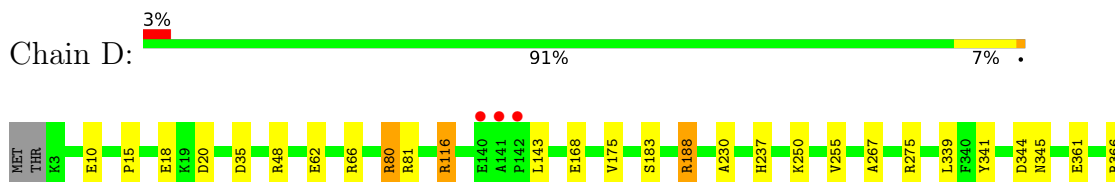
- Molecule 1: Neamine transaminase NeoN

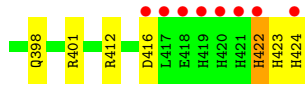


- Molecule 1: Neamine transaminase NeoN



- Molecule 1: Neamine transaminase NeoN





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.73Å 82.70Å 100.86Å 75.59° 88.85° 70.60°	Depositor
Resolution (Å)	97.46 – 1.75 29.52 – 1.75	Depositor EDS
% Data completeness (in resolution range)	95.9 (97.46-1.75) 95.9 (29.52-1.75)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	7.89 (at 1.75Å)	Xtrriage
Refinement program	REFMAC 5.8.0124	Depositor
R, $R_{free}$	0.175 , 0.208 0.188 , 0.219	Depositor DCC
$R_{free}$ test set	8819 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.2	Xtrriage
Anisotropy	0.115	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 43.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13702	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% 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: NA, EDO, PMP

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	0.56	0/3201	0.86	5/4354 (0.1%)
1	B	0.56	0/3290	0.87	10/4474 (0.2%)
1	C	0.54	0/3175	0.83	4/4320 (0.1%)
1	D	0.57	0/3323	0.92	10/4519 (0.2%)
All	All	0.56	0/12989	0.87	29/17667 (0.2%)

There are no bond length outliers.

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	48	ARG	NE-CZ-NH2	10.36	125.48	120.30
1	B	275[A]	ARG	NE-CZ-NH1	-8.43	116.09	120.30
1	B	275[B]	ARG	NE-CZ-NH1	-8.43	116.09	120.30
1	A	116	ARG	NE-CZ-NH1	-8.40	116.10	120.30
1	A	116	ARG	NE-CZ-NH2	7.96	124.28	120.30

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	3120	0	3087	15	0
1	B	3206	0	3141	14	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3103	0	3048	16	0
1	D	3233	0	3160	13	0
2	A	16	0	10	0	0
2	B	16	0	10	0	0
2	C	16	0	10	1	0
2	D	16	0	10	0	0
3	A	4	0	6	0	0
4	C	1	0	0	0	0
5	A	252	0	0	1	0
5	B	240	0	0	2	0
5	C	203	0	0	1	0
5	D	276	0	0	4	0
All	All	13702	0	12482	55	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 2.

The worst 5 of 55 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:B:291:ARG:NH2	1:B:354:GLU:OE2	2.12	0.83
1:B:116:ARG:HG3	1:B:121:ARG:O	1.85	0.77
1:C:161:ARG:NH2	1:C:189:GLU:OE1	2.21	0.73
1:B:36:GLU:OE1	5:B:601:HOH:O	2.07	0.71
1:A:48:ARG:HG3	1:A:341:TYR:CE2	2.32	0.64

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	411/424 (97%)	402 (98%)	8 (2%)	1 (0%)	47 29
1	B	420/424 (99%)	410 (98%)	9 (2%)	1 (0%)	47 29
1	C	409/424 (96%)	401 (98%)	7 (2%)	1 (0%)	47 29
1	D	425/424 (100%)	415 (98%)	9 (2%)	1 (0%)	47 29
All	All	1665/1696 (98%)	1628 (98%)	33 (2%)	4 (0%)	47 29

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	230	ALA
1	B	230	ALA
1	C	230	ALA
1	D	230	ALA

### 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	307/318 (96%)	297 (97%)	10 (3%)	38 15
1	B	315/318 (99%)	304 (96%)	11 (4%)	36 13
1	C	304/318 (96%)	295 (97%)	9 (3%)	41 18
1	D	320/318 (101%)	305 (95%)	15 (5%)	26 7
All	All	1246/1272 (98%)	1201 (96%)	45 (4%)	37 13

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	345	ASN
1	D	143	LEU
1	C	407	THR
1	D	66	ARG
1	D	250	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11

such sidechains are listed below:

Mol	Chain	Res	Type
1	D	237	HIS
1	D	398	GLN
1	D	424	HIS
1	D	423	HIS
1	C	237	HIS

### 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](#)

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	$\# Z  > 2$
3	EDO	A	502	-	3,3,3	0.60	0	2,2,2	0.45	0
2	PMP	C	501	-	16,16,16	2.88	3 (18%)	21,23,23	1.68	6 (28%)
2	PMP	B	501	-	16,16,16	2.45	4 (25%)	21,23,23	2.20	8 (38%)
2	PMP	D	501	-	16,16,16	2.62	3 (18%)	21,23,23	2.37	8 (38%)
2	PMP	A	501	-	16,16,16	2.43	3 (18%)	21,23,23	1.63	6 (28%)

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	502	-	-	0/1/1/1	-
2	PMP	C	501	-	-	2/8/8/8	0/1/1/1
2	PMP	B	501	-	-	0/8/8/8	0/1/1/1
2	PMP	D	501	-	-	2/8/8/8	0/1/1/1
2	PMP	A	501	-	-	3/8/8/8	0/1/1/1

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	501	PMP	C3-C2	8.19	1.49	1.40
2	B	501	PMP	C3-C2	6.98	1.47	1.40
2	D	501	PMP	C3-C2	6.85	1.47	1.40
2	A	501	PMP	C3-C2	6.08	1.47	1.40
2	C	501	PMP	C5-C4	5.63	1.48	1.40

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	501	PMP	C2A-C2-C3	-6.29	113.12	120.89
2	B	501	PMP	C2A-C2-C3	-4.64	115.15	120.89
2	D	501	PMP	C6-C5-C4	4.56	121.34	118.12
2	D	501	PMP	C2A-C2-N1	4.09	125.66	117.67
2	B	501	PMP	C2A-C2-N1	3.45	124.41	117.67

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	501	PMP	C3-C4-C4A-N4A
2	C	501	PMP	C5-C4-C4A-N4A
2	D	501	PMP	C5-C4-C4A-N4A
2	A	501	PMP	C3-C4-C4A-N4A
2	D	501	PMP	C3-C4-C4A-N4A

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	501	PMP	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 [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	409/424 (96%)	-0.29	3 (0%) 87 92	14, 23, 40, 69	0
1	B	419/424 (98%)	-0.11	4 (0%) 82 87	16, 23, 45, 79	0
1	C	410/424 (96%)	-0.08	7 (1%) 70 77	14, 27, 46, 68	0
1	D	422/424 (99%)	-0.13	11 (2%) 56 62	15, 22, 47, 94	0
All	All	1660/1696 (97%)	-0.15	25 (1%) 73 80	14, 24, 45, 94	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	141	ALA	5.8
1	D	422	HIS	5.4
1	D	417	LEU	5.1
1	B	142	PRO	4.9
1	D	142	PRO	4.9

### 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	A	502	4/4	0.91	0.11	34,36,40,42	0
2	PMP	B	501	16/16	0.98	0.09	16,23,26,37	0
2	PMP	C	501	16/16	0.98	0.08	18,23,27,40	0
2	PMP	D	501	16/16	0.98	0.07	16,21,27,42	0
2	PMP	A	501	16/16	0.98	0.10	16,19,24,42	0
4	NA	C	502	1/1	0.99	0.03	20,20,20,20	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.