

# wwPDB X-ray Structure Validation Summary Report (i)

#### Aug 23, 2023 – 03:01 PM EDT

PDB ID : 3BWN

Title: L-tryptophan aminotransferase

Authors: Ferrer, J.-L.; Noel, J.P.; Pojer, F.; Bowman, M.; Chory, J.; Tao, Y.

Deposited on : 2008-01-10

Resolution : 2.25 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
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) 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

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$ 

CCP4 : 7.0.044 (Gargrove) roteins) : Engh & Huber (2001)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

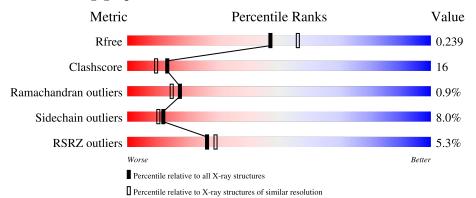
Validation Pipeline (wwPDB-VP) : 2.35

## 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X- $RAY\ DIFFRACTION$ 

The reported resolution of this entry is 2.25 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	Similar resolution
Metric	$(\# { m Entries})$	$(\#  ext{Entries},  ext{ resolution range}(\mathring{A}))$
$R_{free}$	130704	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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 >=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 <=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	391	63%	25%	• 7%				
1	В	391	67%	23%	• 6%				
1	D	391	64%	24%	• 8%				
1	Е	391	57%	27%	8% • 7%				
1	F	391	60%	29%	• • 8%				



Mol	Chain	Length	Quality of chain					
			4%					
2	С	391	64%	24%	•	8%		

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	PMP	A	1001	-	-	X	-
3	PMP	В	1001	-	-	X	-
3	PMP	F	1001[A]	-	-	X	-



## 2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 18137 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 L-tryptophan aminotransferase.

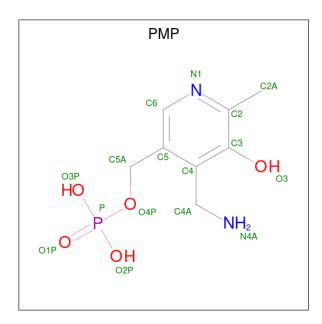
Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
1	A	362	Total	С	N	О	S	0	0	0
1	Λ	302	2906	1848	494	545	19	0	0	
1	В	369	Total	С	N	О	S	0	0	0
1	Ъ	309	2962	1878	505	560	19	0	0	
1	D	359	Total	С	N	О	S	0	1	0
1	ט	309	2886	1833	491	543	19	0		
1	E	362	Total	С	N	О	S	0	0	0
1	<u> 1</u> 2	302	2900	1847	491	543	19	0	0	
1	F	360	Total	С	N	О	S	0	1	0
	Г	300	2892	1840	492	541	19	U	1	U

• Molecule 2 is a protein called L-tryptophan aminotransferase.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace		
2	С	361	Total 2918	C 1855	N 494	O 549	P 1	S 19	0	0	0

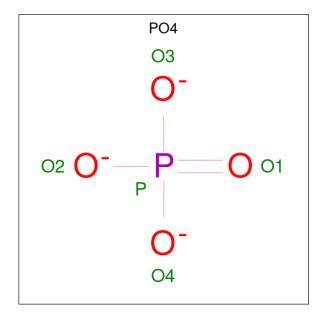
• Molecule 3 is 4'-DEOXY-4'-AMINOPYRIDOXAL-5'-PHOSPHATE (three-letter code: PMP) (formula:  $C_8H_{13}N_2O_5P$ ).





Mol	Chain	Residues	${f Atoms}$			ZeroOcc	AltConf		
3	Λ	1	Total	С	N	О	Р	0	0
3	A	1	16	8	2	5	1	0	U
3	D	1	Total	С	N	О	Р	0	0
)	Ъ	1	16	8	2	5	1	0	U
3	D	1	Total	С	N	О	Р	0	1
)	ע	1	16	8	2	5	1	U	1
3	F	1	Total	С	N	О	Р	0	1
3	Г	1	16	8	2	5	1		1

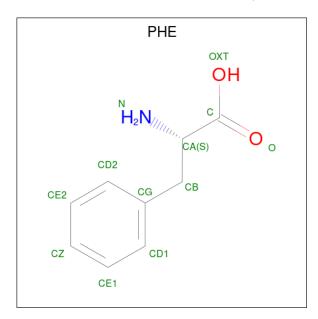
 $\bullet$  Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula:  $\mathrm{O_4P}).$ 





$\mathbf{Mol}$	Chain	Residues	Atoms			ZeroOcc	AltConf
4	Е	1	Total 5	O 4	P 1	0	0

 $\bullet$  Molecule 5 is PHENYLALANINE (three-letter code: PHE) (formula:  $\mathrm{C_9H_{11}NO_2}).$ 



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	F	1	Total 12	C 9	N 1	O 2	0	1

• Molecule 6 is water.

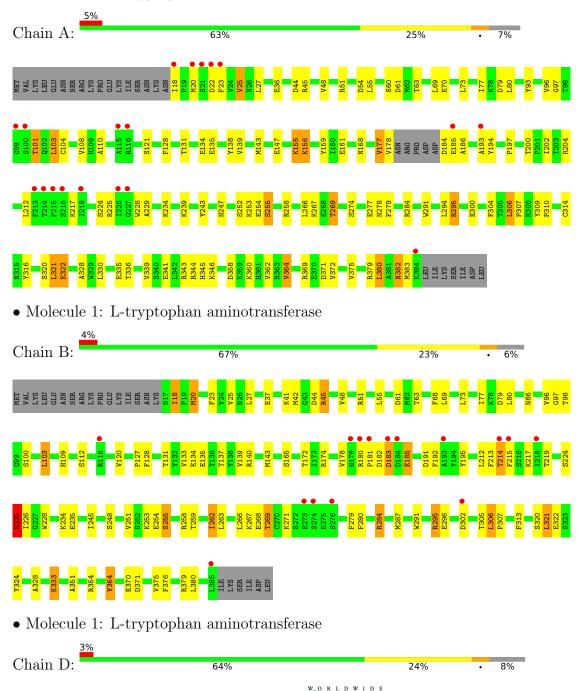
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	102	Total O 102 102	0	0
6	В	129	Total O 129 129	0	0
6	С	90	Total O 90 90	0	0
6	D	100	Total O 100 100	0	0
6	E	75	Total O 75 75	0	0
6	F	96	Total O 96 96	0	0

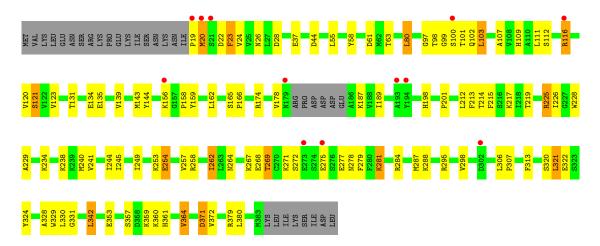


## 3 Residue-property plots (i)

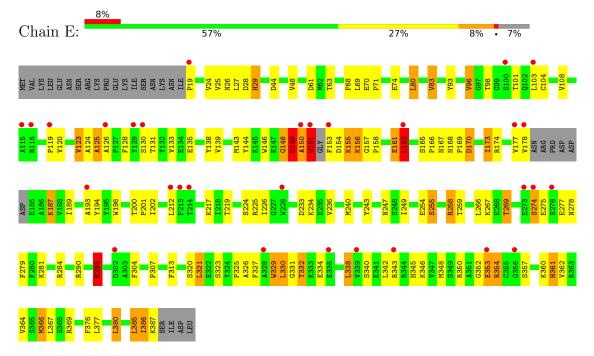
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: L-tryptophan aminotransferase

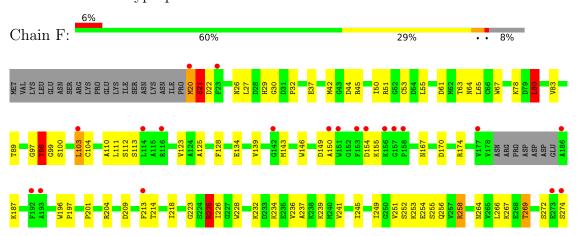




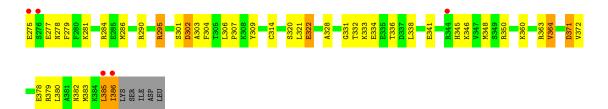
• Molecule 1: L-tryptophan aminotransferase



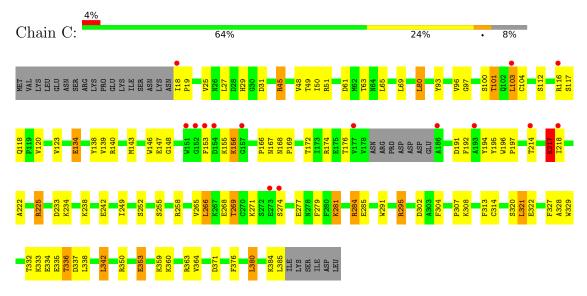
• Molecule 1: L-tryptophan aminotransferase







 $\bullet$  Molecule 2: L-tryptophan aminotransferase





## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	93.41Å 97.83Å 139.46Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $104.35^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	29.66 - 2.25	Depositor
resolution (A)	29.66 - 2.25	EDS
% Data completeness	95.9 (29.66-2.25)	Depositor
(in resolution range)	95.9 (29.66-2.25)	EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.56	Depositor
$< I/\sigma(I) > 1$	2.58  (at  2.24Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
$R, R_{free}$	0.224 , $0.253$	Depositor
it, it free	0.222 , $0.239$	DCC
$R_{free}$ test set	5558 reflections $(5.02%)$	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.9	Xtriage
Anisotropy	0.200	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.32, 45.2	EDS
L-test for twinning <sup>2</sup>	$ < L > = 0.48, < L^2> = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	18137	wwPDB-VP
Average B, all atoms $(Å^2)$	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The analyses of the Patterson function reveals a significant off-origin peak that is 31.12 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1668e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.

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



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

## 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: PO4, PMP, LLP

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).

Mal	Mol Chain		nd lengths	В	ond angles
IVIOI	Chain	RMSZ	# Z  > 5	RMSZ	# Z >5
1	A	0.81	1/2975~(0.0%)	0.84	$2/4021 \ (0.0\%)$
1	В	0.81	0/3033	0.85	4/4104 (0.1%)
1	D	0.83	0/2955	0.88	5/3994 (0.1%)
1	Е	0.75	0/2968	0.88	9/4011 (0.2%)
1	F	0.74	0/2960	0.79	3/4000 (0.1%)
2	С	0.79	0/2962	0.83	3/4003 (0.1%)
All	All	0.79	1/17853 (0.0%)	0.85	$26/24133 \ (0.1\%)$

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 maintain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	0
1	Е	4	8
1	F	0	1
2	С	0	1
All	All	5	10

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(\mathbf{\mathring{A}})$	Ideal(A)
1	A	229	ALA	CA-CB	-5.47	1.41	1.52

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
1	D	225	ARG	NE-CZ-NH2	-11.78	114.41	120.30
1	Е	321	LEU	CA-CB-CG	8.72	135.36	115.30



Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\mathbf{Observed}(^{o})$	$\operatorname{Ideal}({}^o)$
1	Ε	151	TRP	N-CA-C	7.85	132.20	111.00
1	В	321	LEU	CA-CB-CG	7.82	133.29	115.30
1	Е	326	ALA	N-CA-C	-7.38	91.06	111.00

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	185	GLU	CA
1	Е	149	ASP	CA
1	Е	151	TRP	CA
1	Е	155	LYS	CA
1	Е	330	LEU	CA

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	С	217	LLP	Mainchain
1	Е	125	ALA	Peptide
1	Е	148	GLY	Peptide
1	Е	149	ASP	Peptide
1	Е	19	PRO	Peptide

### 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	2906	0	2839	83	0
1	В	2962	0	2880	87	0
1	D	2886	0	2810	95	0
1	Е	2900	0	2832	131	0
1	F	2892	0	2827	101	0
2	С	2918	0	2847	78	0
3	A	16	0	11	13	0
3	В	16	0	11	8	0
3	D	16	0	4	0	0
3	F	16	0	4	6	0
4	Е	5	0	0	1	0



Continued	trom	mmoninonic	maaa
COHABABACA		DIEUIUU	DUIUE
0 0 1000100000			

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	F	12	0	0	0	0
6	A	102	0	0	5	0
6	В	129	0	0	6	0
6	С	90	0	0	9	0
6	D	100	0	0	6	0
6	Е	75	0	0	5	0
6	F	96	0	0	11	0
All	All	18137	0	17065	565	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 16.

The worst 5 of 565 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic	Clash
Atom-1	Atom-2	$\operatorname{distance}\left(\mathrm{\AA}\right)$	overlap (Å)
1:A:217:LYS:HE3	3:A:1001:PMP:N4A	1.36	1.38
1:B:217:LYS:HE3	3:B:1001:PMP:N4A	1.46	1.25
2:C:218:ILE:HD11	6:C:394:HOH:O	1.35	1.23
1:A:217:LYS:CE	3:A:1001:PMP:HNA2	1.61	1.12
1:B:217:LYS:CE	3:B:1001:PMP:HNA2	1.61	1.12

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	358/391 (92%)	330 (92%)	23 (6%)	5 (1%)	11 7
1	В	367/391 (94%)	345 (94%)	20 (5%)	2 (0%)	29 29
1	D	355/391 (91%)	341 (96%)	14 (4%)	0	100 100
1	Е	356/391 (91%)	310 (87%)	39 (11%)	7 (2%)	7 4



Mol	Chain	Analysed	Favoured	Allowed	Outliers	Perce	ntiles
1	F	356/391 (91%)	328 (92%)	22 (6%)	6 (2%)	9	4
2	$\mathbf{C}$	356/391 (91%)	331 (93%)	25 (7%)	0	100	100
All	All	2148/2346 (92%)	1985 (92%)	143 (7%)	20 (1%)	17	14

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	155	LYS
1	В	182	ASP
1	В	183	ASP
1	Е	156	LYS
1	Е	353	GLU

#### 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	315/346 (91%)	290 (92%)	25 (8%)	12	10	
1	В	322/346 (93%)	296 (92%)	26 (8%)	11	10	
1	D	314/346 (91%)	291 (93%)	23 (7%)	14	12	
1	E	314/346 (91%)	283 (90%)	31 (10%)	8	5	
1	F	314/346 (91%)	291 (93%)	23 (7%)	14	12	
2	C	315/345~(91%)	292 (93%)	23 (7%)	14	12	
All	All	1894/2075~(91%)	1743 (92%)	151 (8%)	12	10	

5 of 151 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	Ε	295	ARG
1	F	320	SER
1	Е	338	LEU
1	F	80	LEU
1	F	386	ILE



Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
2	С	374	ASN
2	С	382	ASN
1	F	345	HIS
1	Е	289	ASN
1	F	64	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)

1 non-standard protein/DNA/RNA residue is 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 T	Type	Type Chain	hain Res	les Link	Bond lengths			Bond angles		
	туре	Chain	nes	Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ   #  Z  > 2	
2	LLP	С	217	2	23,24,25	2.25	5 (21%)	25,32,34	1.39	4 (16%)

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
2	LLP	С	217	2	-	8/16/17/19	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
2	С	217	LLP	C4'-NZ	6.00	1.47	1.27
2	С	217	LLP	C4-C5	5.51	1.48	1.42
2	С	217	LLP	C6-C5	4.37	1.46	1.37



Mol	Chain	Res	Type	Atoms	Z	Observed(A)	$Ideal(\AA)$
2	С	217	LLP	C4-C4'	2.75	1.51	1.46
2	С	217	LLP	CB-CA	-2.16	1.50	1.53

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
2	С	217	LLP	C4-C4'-NZ	-3.27	109.29	124.31
2	С	217	LLP	C5'-C5-C6	-3.20	114.11	119.37
2	С	217	LLP	C3-C4-C5	-2.64	116.24	118.26
2	С	217	LLP	C5-C6-N1	-2.49	119.67	123.82

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	С	217	LLP	C5'-OP4-P-OP1
2	С	217	LLP	C5'-OP4-P-OP2
2	С	217	LLP	C5'-OP4-P-OP3
2	С	217	LLP	C4-C4'-NZ-CE
2	С	217	LLP	C3-C4-C4'-NZ

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	$\operatorname{Res}$	Type	Clashes	Symm-Clashes
2	С	217	LLP	5	0

### 5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry (i)

6 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	es Link	Bond lengths			Bond angles		
MIOI	туре	Chain	nes	Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	PMP	В	1001	-	16,16,16	2.26	3 (18%)	21,23,23	2.15	8 (38%)
3	PMP	A	1001	-	16,16,16	2.28	4 (25%)	21,23,23	1.56	5 (23%)
4	PO4	Е	1001	-	4,4,4	0.87	0	6,6,6	0.48	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	PMP	В	1001	-	-	1/8/8/8	0/1/1/1
3	PMP	A	1001	-	-	2/8/8/8	0/1/1/1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$\operatorname{Observed}(\text{\AA})$	Ideal(A)
3	В	1001	PMP	C5-C4	5.66	1.48	1.40
3	A	1001	PMP	C5-C4	5.15	1.47	1.40
3	В	1001	PMP	C3-C4	4.90	1.47	1.40
3	A	1001	PMP	C6-C5	4.88	1.47	1.37
3	В	1001	PMP	C6-C5	4.42	1.47	1.37

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

Mol	Chain	Res	Type	Atoms	$\mathbf{Z}$	$Observed(^o)$	$Ideal(^{o})$
3	В	1001	PMP	C6-C5-C4	-6.02	113.86	118.12
3	A	1001	PMP	O2P-P-O4P	-3.71	96.85	106.73
3	A	1001	PMP	O3-C3-C2	3.46	125.03	117.49
3	В	1001	PMP	C5A-C5-C6	-3.17	114.15	119.37
3	В	1001	PMP	C6-N1-C2	3.03	124.79	119.17

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1001	PMP	C5-C4-C4A-N4A
3	A	1001	PMP	C3-C4-C4A-N4A
3	В	1001	PMP	C4-C5-C5A-O4P



There are no ring outliers.

3 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	В	1001	PMP	8	0
3	A	1001	PMP	13	0
4	Е	1001	PO4	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,  $95^{th}$  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>	$\# \mathrm{RSRZ}{>}2$	$OWAB(A^2)$	Q<0.9
1	A	362/391~(92%)	0.14	19 (5%) 27 30	24, 45, 70, 88	0
1	В	369/391 (94%)	0.01	15 (4%) 37 40	21, 40, 69, 82	0
1	D	359/391 (91%)	-0.02	12 (3%) 46 48	22, 40, 67, 91	0
1	E	$362/391 \ (92\%)$	0.47	30 (8%) 11 12	29, 54, 77, 86	0
1	F	360/391 (92%)	0.28	25 (6%) 16 18	26, 49, 75, 88	0
2	С	360/391 (92%)	0.12	15 (4%) 36 38	23, 44, 66, 76	0
All	All	2172/2346 (92%)	0.17	116 (5%) 26 29	21, 45, 72, 91	0

The worst 5 of 116 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	385	LEU	10.2
1	F	385	LEU	9.0
1	Е	335	GLU	6.0
1	Е	151	TRP	5.9
1	Е	150	ALA	5.0

### 6.2 Non-standard residues in protein, DNA, RNA chains (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,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
2	LLP	С	217	24/25	0.95	0.21	41,53,55,58	0



#### 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,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
5	PHE	F	2001[B]	12/12	0.88	0.26	29,31,33,35	0
4	PO4	Ε	1001	5/5	0.94	0.20	61,61,62,62	0
3	PMP	D	1001[A]	16/16	0.94	0.22	48,50,54,57	0
3	PMP	F	1001[A]	16/16	0.95	0.18	55,57,59,60	0
3	PMP	В	1001	16/16	0.95	0.15	54,55,59,60	0
3	PMP	A	1001	16/16	0.95	0.20	54,57,60,60	0

#### 6.5 Other polymers (i)

There are no such residues in this entry.

