WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 31, 2023 – 07:42 PM JST

PDB ID	:	5GZU
Title	:	Crystal Structure of Chitinase ChiW from Paenibacillus sp. str. FPU-7 Reveals a Novel Type of Bacterial Cell-Surface-Expressed Multi-Modular Enzyme Machinery
Authors	:	Itoh, T.; Hibi, T.; Suzuki, F.; Sugimoto, I.; Fujiwara, A.; Inaka, K.; Tanaka, H.; Ohta, K.; Fujii, Y.; Taketo, A.; Kimoto, H.
Deposited on	:	2016-10-01
Resolution	:	2.03 Å(reported)

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

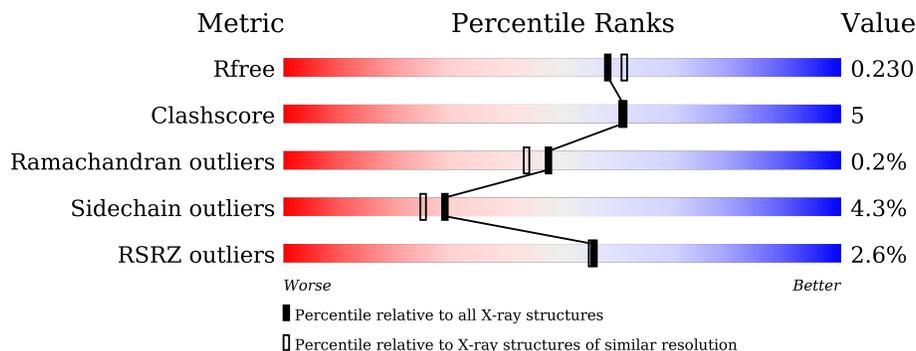|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.03 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	Percentile Ranks	Value
Rfree		0.230
Clashscore		5
Ramachandran outliers		0.2%
Sidechain outliers		4.3%
RSRZ outliers		2.6%

Worse ——— Better

▌ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	10434 (2.04-2.00)
Clashscore	141614	11643 (2.04-2.00)
Ramachandran outliers	138981	11493 (2.04-2.00)
Sidechain outliers	138945	11492 (2.04-2.00)
RSRZ outliers	127900	10220 (2.04-2.00)

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	885	2% / 81% / 14% · ·
1	B	885	3% / 84% / 12% · ·

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 14861 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 Chitinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	858	Total 6753	C 4318	N 1109	O 1315	S 11	0	5	0
1	B	859	Total 6776	C 4337	N 1111	O 1314	S 14	0	7	0

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	534	MET	-	initiating methionine	UNP K7ZLW6
A	535	ASN	-	expression tag	UNP K7ZLW6
A	536	HIS	-	expression tag	UNP K7ZLW6
A	537	LYS	-	expression tag	UNP K7ZLW6
A	538	VAL	-	expression tag	UNP K7ZLW6
A	539	HIS	-	expression tag	UNP K7ZLW6
A	540	HIS	-	expression tag	UNP K7ZLW6
A	541	HIS	-	expression tag	UNP K7ZLW6
A	542	HIS	-	expression tag	UNP K7ZLW6
A	543	HIS	-	expression tag	UNP K7ZLW6
A	544	HIS	-	expression tag	UNP K7ZLW6
A	545	ILE	-	expression tag	UNP K7ZLW6
A	546	GLU	-	expression tag	UNP K7ZLW6
A	547	GLY	-	expression tag	UNP K7ZLW6
A	548	ARG	-	expression tag	UNP K7ZLW6
A	549	HIS	-	expression tag	UNP K7ZLW6
A	550	MET	-	expression tag	UNP K7ZLW6
A	551	GLU	-	expression tag	UNP K7ZLW6
A	552	LEU	-	expression tag	UNP K7ZLW6
A	553	GLY	-	expression tag	UNP K7ZLW6
A	554	THR	-	expression tag	UNP K7ZLW6
A	555	LEU	-	expression tag	UNP K7ZLW6
A	556	GLU	-	expression tag	UNP K7ZLW6
B	534	MET	-	initiating methionine	UNP K7ZLW6
B	535	ASN	-	expression tag	UNP K7ZLW6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	536	HIS	-	expression tag	UNP K7ZLW6
B	537	LYS	-	expression tag	UNP K7ZLW6
B	538	VAL	-	expression tag	UNP K7ZLW6
B	539	HIS	-	expression tag	UNP K7ZLW6
B	540	HIS	-	expression tag	UNP K7ZLW6
B	541	HIS	-	expression tag	UNP K7ZLW6
B	542	HIS	-	expression tag	UNP K7ZLW6
B	543	HIS	-	expression tag	UNP K7ZLW6
B	544	HIS	-	expression tag	UNP K7ZLW6
B	545	ILE	-	expression tag	UNP K7ZLW6
B	546	GLU	-	expression tag	UNP K7ZLW6
B	547	GLY	-	expression tag	UNP K7ZLW6
B	548	ARG	-	expression tag	UNP K7ZLW6
B	549	HIS	-	expression tag	UNP K7ZLW6
B	550	MET	-	expression tag	UNP K7ZLW6
B	551	GLU	-	expression tag	UNP K7ZLW6
B	552	LEU	-	expression tag	UNP K7ZLW6
B	553	GLY	-	expression tag	UNP K7ZLW6
B	554	THR	-	expression tag	UNP K7ZLW6
B	555	LEU	-	expression tag	UNP K7ZLW6
B	556	GLU	-	expression tag	UNP K7ZLW6

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).
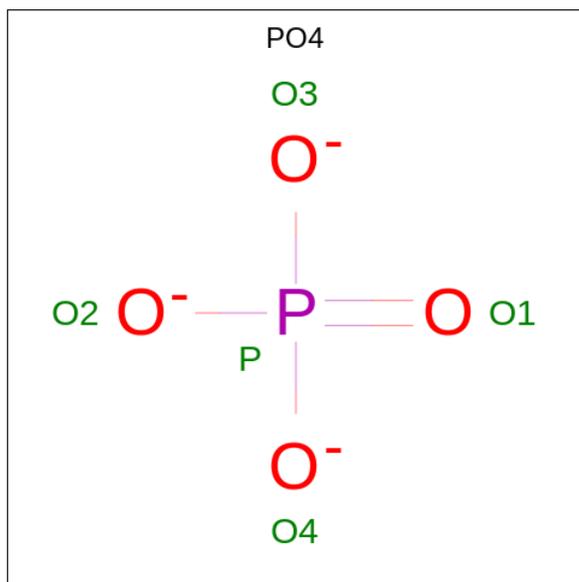
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 5	O 4	P 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total 5	O 4	P 1	0	0

- Molecule 3 is water.

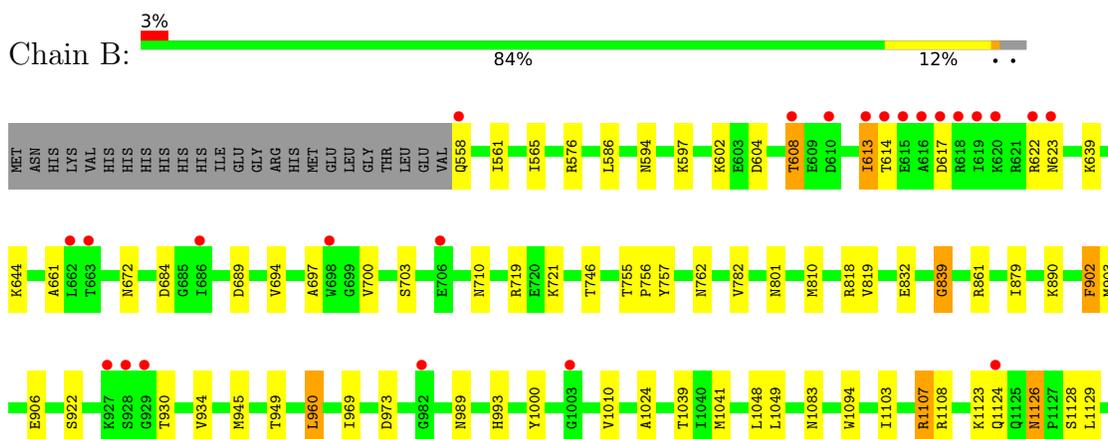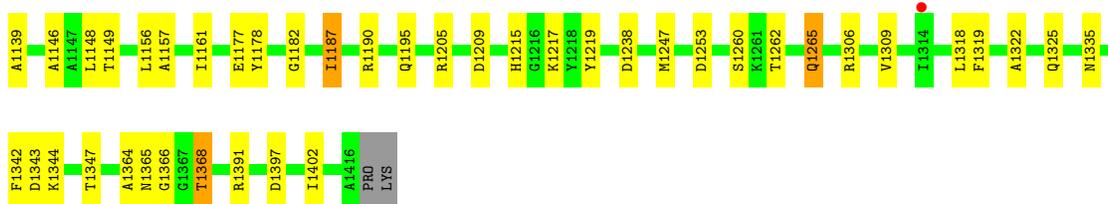| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	652	Total 652	O 652	0	0
3	B	670	Total 670	O 670	0	0

3 Residue-property plots ⓘ

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: Chitinase

- Molecule 1: Chitinase

A1139 A1146 A1147 L1148 T1149 L1156 A1157 I1161 E1177 Y1178 G1182 I1187 R1190 Q1195 R1205 D1209 H1215 G1216 K1217 Y1218 Y1219 D1238 M1247 D1253 S1260 K1261 T1262 Q1265 R1306 V1309 I1314 L1318 F1319 A1322 Q1325 N1335

F1342 D1343 K1344 T1347 A1364 N1365 G1366 G1367 T1368 R1391 D1397 I1402 A1416 PRO LYS
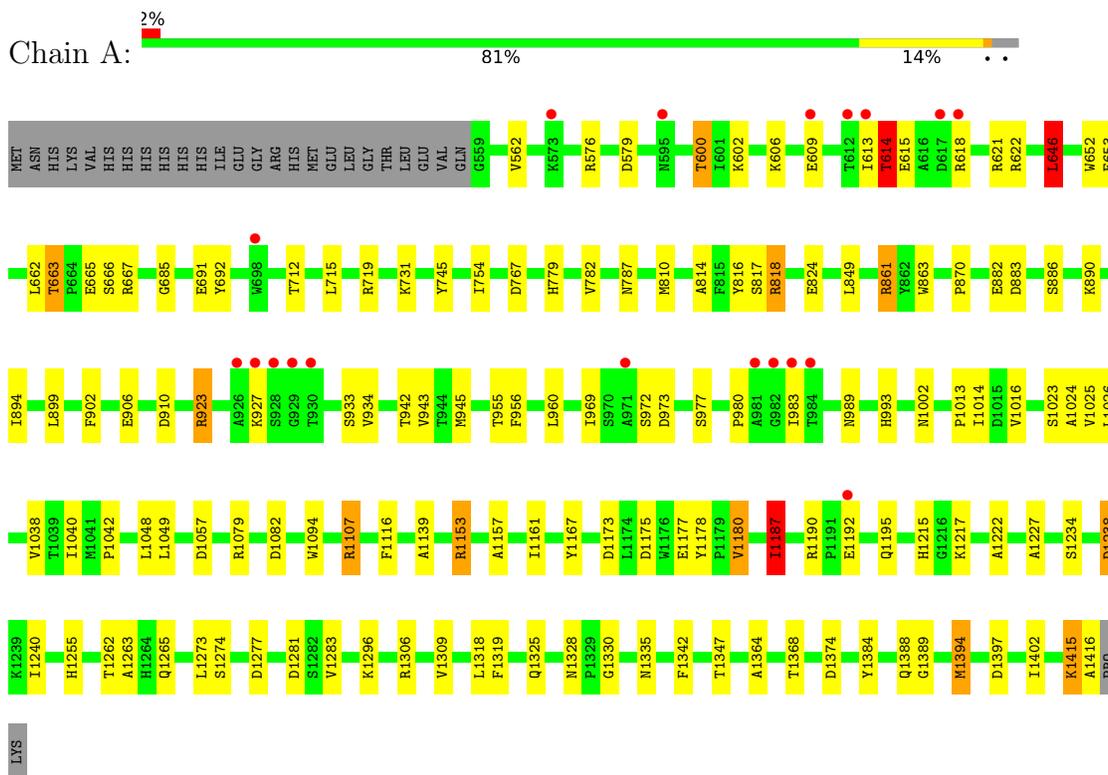
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	111.98Å 127.99Å 162.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.01 – 2.03 41.28 – 2.03	Depositor EDS
% Data completeness (in resolution range)	98.8 (50.01-2.03) 98.8 (41.28-2.03)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.27 (at 2.03Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, R_{free}	0.182 , 0.226 0.190 , 0.230	Depositor DCC
R_{free} test set	7476 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	24.0	Xtriage
Anisotropy	0.029	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 50.1	EDS
L-test for twinning[2]	$< |L| > = 0.43, < L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	14861	wwPDB-VP
Average B, all atoms (Å^2)	31.0	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 22.42 % 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 5.7242e-03.*

---

[1]Intensities estimated from amplitudes.
[2]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.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	1.02	3/6942 (0.0%)	1.09	36/9460 (0.4%)
1	B	1.01	6/6969 (0.1%)	1.06	27/9494 (0.3%)
All	All	1.01	9/13911 (0.1%)	1.07	63/18954 (0.3%)

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	B	0	1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	839	GLY	N-CA	6.54	1.55	1.46
1	A	745	TYR	CE1-CZ	-6.49	1.30	1.38
1	A	1306	ARG	CD-NE	-6.17	1.35	1.46
1	B	839	GLY	CA-C	5.84	1.61	1.51
1	B	1306	ARG	CD-NE	-5.78	1.36	1.46

The worst 5 of 63 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1306	ARG	NE-CZ-NH2	-19.03	110.78	120.30
1	B	1306	ARG	NE-CZ-NH2	-18.56	111.02	120.30
1	B	1306	ARG	NE-CZ-NH1	16.26	128.43	120.30
1	A	1306	ARG	NE-CZ-NH1	14.16	127.38	120.30
1	A	818	ARG	NE-CZ-NH1	12.81	126.70	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	839	GLY	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	6753	0	6529	73	0
1	B	6776	0	6561	48	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
3	A	652	0	0	7	0
3	B	670	0	0	10	0
All	All	14861	0	13090	121	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 121 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:A:1335:ASN:HD21	1:A:1402:ILE:H	1.21	0.86
1:B:694:VAL:O	1:B:703:SER:OG	1.94	0.85
1:A:614:THR:HG22	1:A:618:ARG:HE	1.47	0.79
1:B:1335:ASN:HD21	1:B:1402:ILE:H	1.29	0.78
1:A:1002:ASN:HB2	3:A:3552:HOH:O	1.84	0.76

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	861/885 (97%)	828 (96%)	31 (4%)	2 (0%)	47	43
1	B	864/885 (98%)	835 (97%)	27 (3%)	2 (0%)	47	43
All	All	1725/1770 (98%)	1663 (96%)	58 (3%)	4 (0%)	47	43

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	614	THR
1	A	615	GLU
1	B	613	ILE
1	B	1365	ASN

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	710/730 (97%)	678 (96%)	32 (4%)	27	23
1	B	712/730 (98%)	681 (96%)	31 (4%)	28	24
All	All	1422/1460 (97%)	1359 (96%)	63 (4%)	29	24

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1273	LEU
1	B	1126	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	597	LYS
1	B	1123[B]	LYS
1	B	1318	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1321	GLN
1	B	1325	GLN
1	B	1365	ASN
1	B	558	GLN
1	A	1335	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 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	$\#|Z| > 2$
2	PO4	B	1501	-	4,4,4	1.26	0	6,6,6	1.34	1 (16%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z|>2$	Counts	RMSZ	$\#|Z|>2$
2	PO4	A	3001	-	4,4,4	1.39	0	6,6,6	1.53	1 (16%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	3001	PO4	O2-P-O1	-2.91	100.23	110.89
2	B	1501	PO4	O4-P-O3	-2.90	98.68	107.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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, 95th 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^2)	Q<0.9
1	A	858/885 (96%)	-0.04	19 (2%)	62	61	14, 29, 54, 106	0
1	B	859/885 (97%)	-0.06	25 (2%)	51	51	12, 26, 52, 131	0
All	All	1717/1770 (97%)	-0.05	44 (2%)	56	55	12, 28, 53, 131	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	614	THR	9.4
1	A	929	GLY	8.4
1	B	619	ILE	6.6
1	B	929	GLY	6.1
1	A	928	SER	5.6

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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, 95th 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(Å²)	Q<0.9
2	PO4	A	3001	5/5	0.99	0.10	25,26,28,29	0
2	PO4	B	1501	5/5	0.99	0.07	23,24,26,29	0

6.5　Other polymers ⓘ

There are no such residues in this entry.