



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 28, 2016 – 01:39 AM EST

PDB ID : 5GZT
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 : 2.10 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

<http://wwpdb.org/validation/2016/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	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

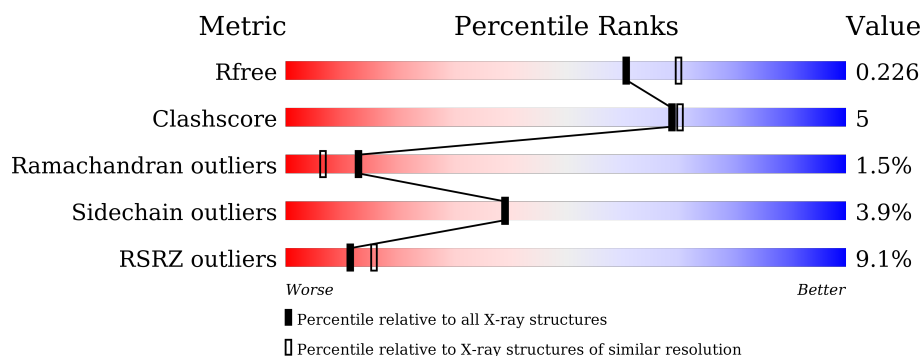
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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}	91344	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	108	<div> <div>34%</div> <div>43%</div> <div>18%</div> <div>••</div> <div>36%</div> </div>
2	B	1136	<div> <div>6%</div> <div>86%</div> <div>12%</div> <div>•</div> </div>

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
5	FMT	B	4003	-	-	-	X
5	FMT	B	4004	-	-	-	X
5	FMT	B	4005	-	-	-	X
5	FMT	B	4006	-	-	X	-
5	FMT	B	4007	-	-	-	X
5	FMT	B	4010	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10611 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	69	Total	C	N	O	0	0	0
			481	297	81	103			

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	175	MET	-	initiating methionine	UNP K7ZLW6
A	176	ASN	-	expression tag	UNP K7ZLW6
A	177	HIS	-	expression tag	UNP K7ZLW6
A	178	LYS	-	expression tag	UNP K7ZLW6
A	179	VAL	-	expression tag	UNP K7ZLW6
A	180	HIS	-	expression tag	UNP K7ZLW6
A	181	HIS	-	expression tag	UNP K7ZLW6
A	182	HIS	-	expression tag	UNP K7ZLW6
A	183	HIS	-	expression tag	UNP K7ZLW6
A	184	HIS	-	expression tag	UNP K7ZLW6
A	185	HIS	-	expression tag	UNP K7ZLW6
A	186	ILE	-	expression tag	UNP K7ZLW6
A	187	GLU	-	expression tag	UNP K7ZLW6
A	188	GLY	-	expression tag	UNP K7ZLW6
A	189	ARG	-	expression tag	UNP K7ZLW6
A	190	HIS	-	expression tag	UNP K7ZLW6
A	191	MET	-	expression tag	UNP K7ZLW6
A	192	GLU	-	expression tag	UNP K7ZLW6
A	193	LEU	-	expression tag	UNP K7ZLW6
A	194	GLY	-	expression tag	UNP K7ZLW6
A	195	THR	-	expression tag	UNP K7ZLW6
A	196	LEU	-	expression tag	UNP K7ZLW6
A	197	GLU	-	expression tag	UNP K7ZLW6

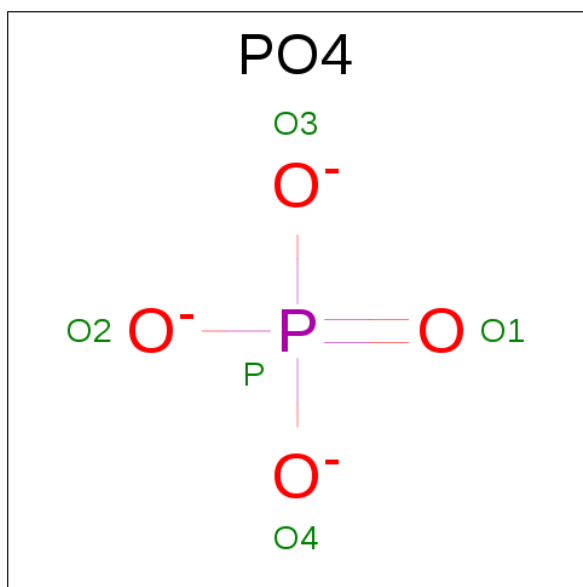
- Molecule 2 is a protein called Chitinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1134	Total	C	N	O	S	0	6	0
			8736	5560	1443	1721	12			

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

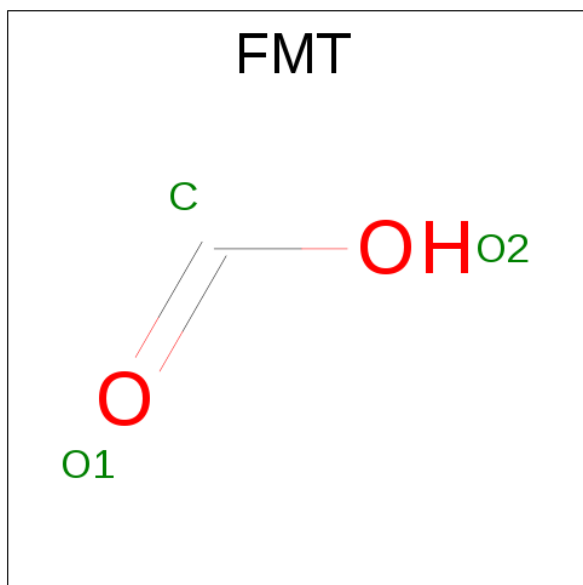
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Na	0	0
			1	1		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



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

- Molecule 5 is FORMIC ACID (three-letter code: FMT) (formula: CH₂O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			3	1	2		
5	B	1	Total	C	O	0	0
			3	1	2		
5	B	1	Total	C	O	0	0
			3	1	2		
5	B	1	Total	C	O	0	0
			3	1	2		
5	B	1	Total	C	O	0	0
			3	1	2		
5	B	1	Total	C	O	0	0
			3	1	2		
5	B	1	Total	C	O	0	0
			3	1	2		
5	B	1	Total	C	O	0	0
			3	1	2		
5	B	1	Total	C	O	0	0
			3	1	2		

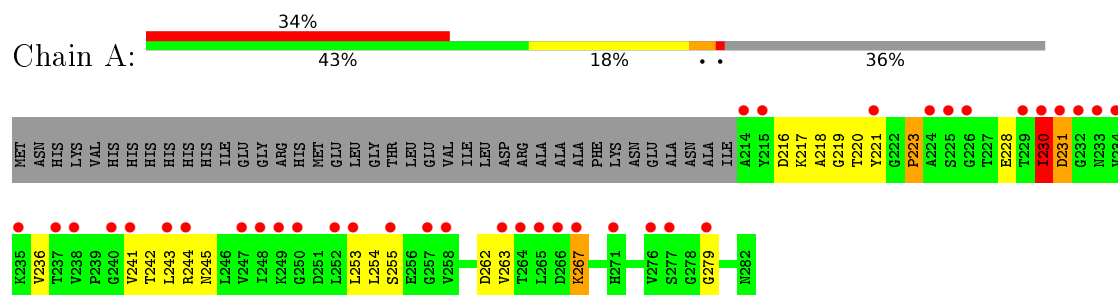
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	18	Total	O	0	0
			18	18		
6	B	1340	Total	O	0	0
			1340	1340		

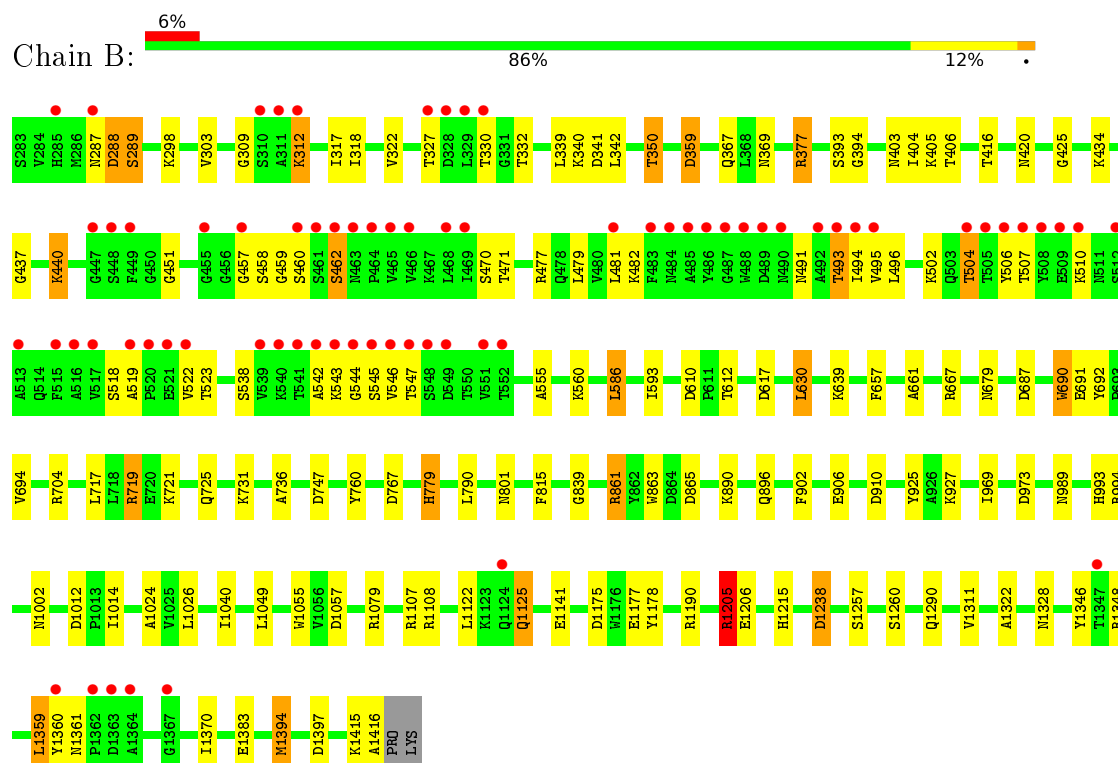
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 2: Chitinase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	114.34Å 123.49Å 130.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.01 – 2.10 48.23 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.01-2.10) 99.7 (48.23-2.10)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.31 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, R_{free}	0.175 , 0.221 0.185 , 0.226	Depositor DCC
R_{free} test set	5370 reflections (5.23%)	DCC
Wilson B-factor (Å ²)	24.6	Xtriage
Anisotropy	0.064	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 58.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	10611	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.97% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA, FMT, 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	0.88	0/485	0.97	0/659
2	B	1.05	7/8954 (0.1%)	1.06	35/12192 (0.3%)
All	All	1.04	7/9439 (0.1%)	1.06	35/12851 (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	A	0	2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1141	GLU	CD-OE2	8.05	1.34	1.25
2	B	1206	GLU	CD-OE2	-6.90	1.18	1.25
2	B	1205	ARG	CD-NE	-6.89	1.34	1.46
2	B	925	TYR	CE1-CZ	5.64	1.45	1.38
2	B	394	GLY	N-CA	5.32	1.54	1.46

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1205	ARG	NE-CZ-NH1	27.74	134.17	120.30
2	B	1205	ARG	NE-CZ-NH2	-22.51	109.05	120.30
2	B	1205	ARG	CD-NE-CZ	7.93	134.70	123.60
2	B	1238[A]	ASP	CB-CG-OD1	7.83	125.35	118.30
2	B	1238[C]	ASP	CB-CG-OD1	7.83	125.35	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	218	ALA	Peptide
1	A	230	ILE	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	481	0	483	9	0
2	B	8736	0	8553	76	0
3	B	1	0	0	0	0
4	B	5	0	0	0	0
5	B	30	0	14	4	0
6	A	18	0	0	1	0
6	B	1340	0	0	14	1
All	All	10611	0	9050	85	1

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 85 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)
2:B:767:ASP:H	2:B:779:HIS:HD2	1.23	0.86
2:B:495:VAL:HG12	2:B:538:SER:OG	1.76	0.86
2:B:312:LYS:NZ	2:B:332:THR:OG1	2.09	0.85
2:B:993:HIS:HD2	6:B:5115:HOH:O	1.62	0.82
2:B:679:ASN:HD21	2:B:725:GLN:HE22	1.26	0.81

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:B:4775:HOH:O	6:B:4968:HOH:O[4_555]	1.57	0.63

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	67/108 (62%)	51 (76%)	8 (12%)	8 (12%)	0	0
2	B	1138/1136 (100%)	1087 (96%)	41 (4%)	10 (1%)	21	15
All	All	1205/1244 (97%)	1138 (94%)	49 (4%)	18 (2%)	13	7

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	255	SER
2	B	288	ASP
2	B	460	SER
2	B	462	SER
2	B	519	ALA

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	53/85 (62%)	47 (89%)	6 (11%)	7	4
2	B	928/924 (100%)	896 (97%)	32 (3%)	44	45
All	All	981/1009 (97%)	943 (96%)	38 (4%)	39	39

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	510	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	547	THR
2	B	1125	GLN
2	B	522	VAL
2	B	586	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	679	ASN
2	B	779	HIS
2	B	1125	GLN
2	B	484	ASN
2	B	993	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 1 is monoatomic - leaving 11 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
4	PO4	B	4002	-	4,4,4	1.59	1 (25%)	6,6,6	0.36	0
5	FMT	B	4003	-	0,2,2	0.00	-	0,1,1	0.00	-
5	FMT	B	4004	-	0,2,2	0.00	-	0,1,1	0.00	-
5	FMT	B	4005	-	0,2,2	0.00	-	0,1,1	0.00	-
5	FMT	B	4006	-	0,2,2	0.00	-	0,1,1	0.00	-
5	FMT	B	4007	-	0,2,2	0.00	-	0,1,1	0.00	-
5	FMT	B	4008	-	0,2,2	0.00	-	0,1,1	0.00	-
5	FMT	B	4009	-	0,2,2	0.00	-	0,1,1	0.00	-
5	FMT	B	4010	-	0,2,2	0.00	-	0,1,1	0.00	-
5	FMT	B	4011	-	0,2,2	0.00	-	0,1,1	0.00	-
5	FMT	B	4012	-	0,2,2	0.00	-	0,1,1	0.00	-

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	PO4	B	4002	-	-	0/0/0/0	0/0/0/0
5	FMT	B	4003	-	-	0/0/0/0	0/0/0/0
5	FMT	B	4004	-	-	0/0/0/0	0/0/0/0
5	FMT	B	4005	-	-	0/0/0/0	0/0/0/0
5	FMT	B	4006	-	-	0/0/0/0	0/0/0/0
5	FMT	B	4007	-	-	0/0/0/0	0/0/0/0
5	FMT	B	4008	-	-	0/0/0/0	0/0/0/0
5	FMT	B	4009	-	-	0/0/0/0	0/0/0/0
5	FMT	B	4010	-	-	0/0/0/0	0/0/0/0
5	FMT	B	4011	-	-	0/0/0/0	0/0/0/0
5	FMT	B	4012	-	-	0/0/0/0	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	4002	PO4	P-O3	-2.61	1.45	1.53

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	4006	FMT	3	0
5	B	4007	FMT	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, 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	69/108 (63%)	2.48	37 (53%) 0 0	59, 83, 113, 119	0
2	B	1134/1136 (99%)	-0.01	72 (6%) 23 31	12, 29, 76, 123	0
All	All	1203/1244 (96%)	0.13	109 (9%) 11 16	12, 31, 86, 123	0

The worst 5 of 109 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	542	ALA	7.7
1	A	221	TYR	7.3
1	A	233	ASN	7.1
1	A	232	GLY	7.1
1	A	215	TYR	6.7

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
5	FMT	B	4005	3/3	0.72	0.33	8.39	49,49,56,61	0
5	FMT	B	4004	3/3	0.68	0.34	7.05	57,57,58,69	0
5	FMT	B	4003	3/3	0.86	0.29	5.07	41,41,45,50	0
5	FMT	B	4010	3/3	0.96	0.13	3.27	37,37,38,41	0
5	FMT	B	4007	3/3	0.94	0.15	2.50	25,25,37,39	0
5	FMT	B	4012	3/3	0.85	0.13	1.85	47,47,52,57	0
5	FMT	B	4006	3/3	0.93	0.15	1.23	25,25,26,39	0
3	NA	B	4001	1/1	0.82	0.20	0.78	50,50,50,50	0
5	FMT	B	4011	3/3	0.98	0.10	-0.71	28,28,29,30	0
4	PO4	B	4002	5/5	1.00	0.08	-1.61	16,17,18,20	0
5	FMT	B	4008	3/3	0.91	0.20	-	33,33,37,44	0
5	FMT	B	4009	3/3	0.84	0.32	-	51,51,56,67	0

6.5 Other polymers [i](#)

There are no such residues in this entry.