wwPDB X-ray Structure Validation Summary Report (i)

Jan 7, 2024 – 07:59 pm GMT

PDB ID	:	5O6K
Title	:	Structure of Polyphosphate Kinase from Meiothermus ruber N121D
Authors	:	Kemper, F.; Gerhardt, S.; Einsle, O.
Deposited on	:	2017-06-06
Resolution	:	2.90 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 (i)) were used in the production of this report:

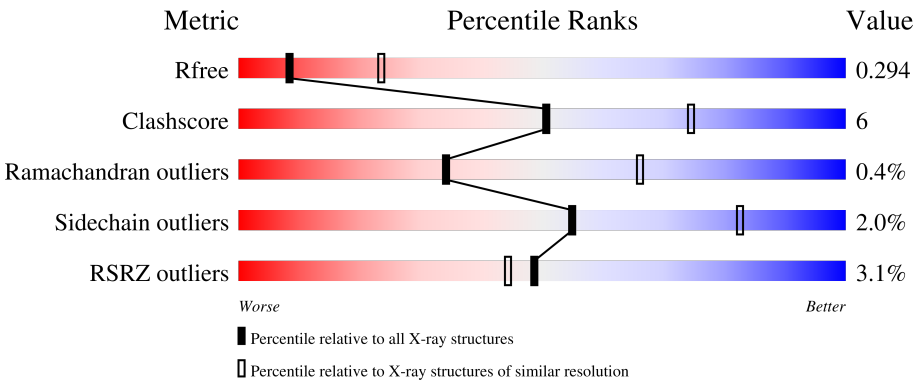
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	267	84% 13% •
1	B	267	% 79% 16% •
1	C	267	4% 78% 18% • •
1	D	267	7% 80% 17% •

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
2	PO4	A	304	-	-	X	-
2	PO4	D	302	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8740 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.

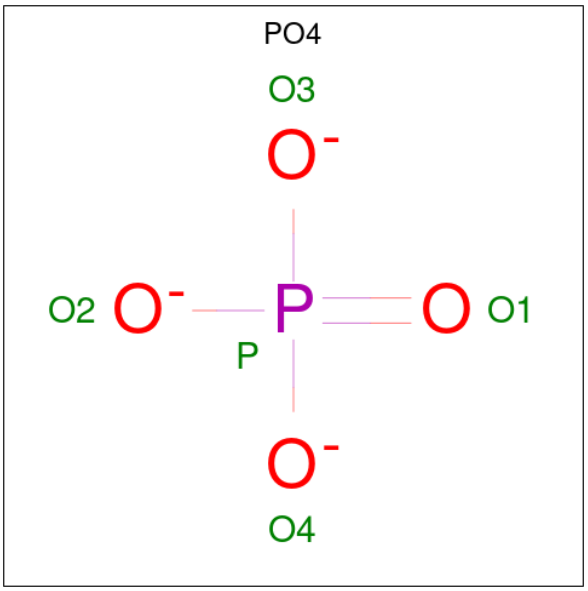
- Molecule 1 is a protein called Polyphosphate:AMP phosphotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	258	2163	1376	394	388	5	0	0	0
1	B	256	2147	1366	392	384	5	0	0	0
1	C	260	2181	1391	394	392	4	0	0	0
1	D	260	2175	1388	394	388	5	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

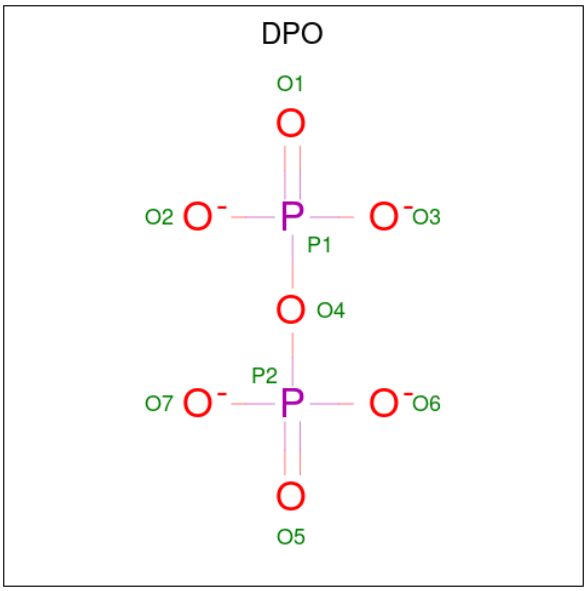
Chain	Residue	Modelled	Actual	Comment	Reference
A	121	ASP	ASN	engineered mutation	UNP A0A0S7ASE9
B	121	ASP	ASN	engineered mutation	UNP A0A0S7ASE9
C	121	ASP	ASN	engineered mutation	UNP A0A0S7ASE9
D	121	ASP	ASN	engineered mutation	UNP A0A0S7ASE9

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



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

- Molecule 3 is DIPHOSPHATE (three-letter code: DPO) (formula: O_7P_2).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total O P 9 7 2	0	0
3	B	1	Total O P 9 7 2	0	0

- Molecule 4 is water.

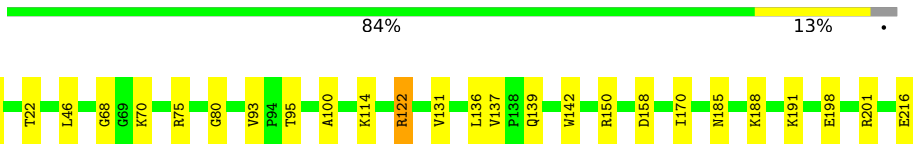
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	3	Total O 3 3	0	0
4	B	2	Total O 2 2	0	0
4	D	1	Total O 1 1	0	0

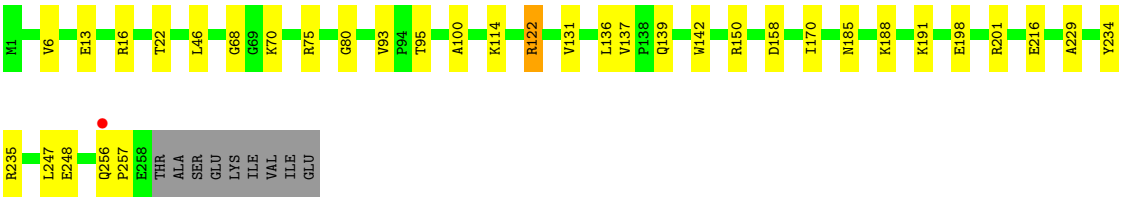
3 Residue-property plots (i)

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: Polyphosphate:AMP phosphotransferase

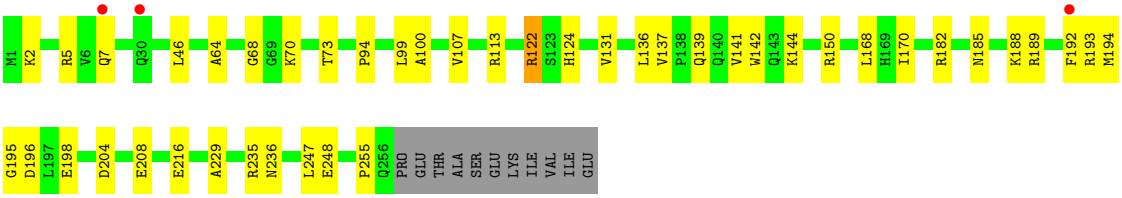
Chain A:





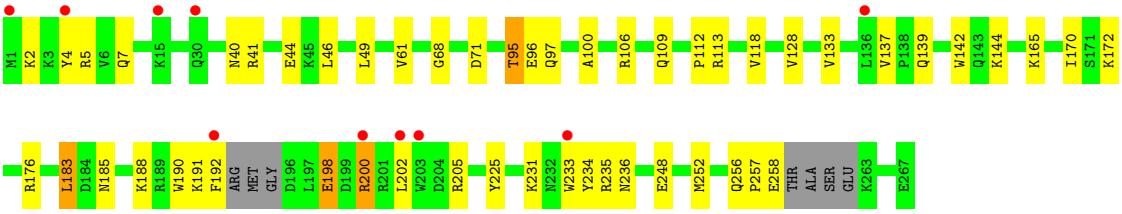
• Molecule 1: Polyphosphate:AMP phosphotransferase

Chain B:



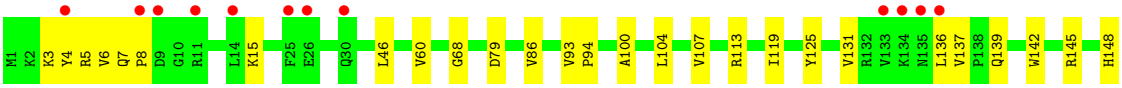
• Molecule 1: Polyphosphate:AMP phosphotransferase

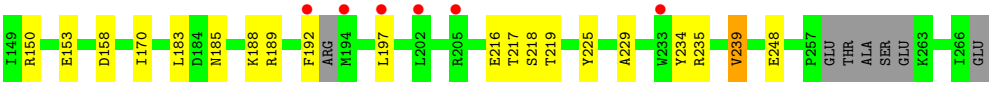
Chain C:



• Molecule 1: Polyphosphate:AMP phosphotransferase

Chain D:




I149
R150
E153
D158
I170
L183
D184
N185
K188
R189
F192
ARG
M194
L197
L202
R205
E216
T217
S218
T219
Y225
A229
W233
Y234
R235
V239
E248
P257
GLU
THR
ALA
SER
GLU
K263
I266
GLU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	166.55Å 166.55Å 94.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.91 – 2.90 47.91 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.6 (47.91-2.90) 99.9 (47.91-2.90)	Depositor EDS
R_{merge}	0.32	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.66 (at 2.91Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.260 , 0.292 0.262 , 0.294	Depositor DCC
R_{free} test set	1507 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	51.0	Xtriage
Anisotropy	0.828	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 10.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	8740	wwPDB-VP
Average B, all atoms (Å^2)	66.0	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 25.09 % 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 3.3871e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: DPO, 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	0.25	0/2214	0.39	0/2991
1	B	0.29	1/2197 (0.0%)	0.39	0/2967
1	C	0.29	1/2230 (0.0%)	0.41	0/3011
1	D	0.27	1/2224 (0.0%)	0.43	0/3002
All	All	0.27	3/8865 (0.0%)	0.41	0/11971

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	7	GLN	C-N	8.21	1.49	1.34
1	B	7	GLN	C-N	7.68	1.48	1.34
1	D	7	GLN	C-N	5.61	1.45	1.34

There are no bond angle outliers.

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	2163	0	2143	21	1
1	B	2147	0	2130	28	0
1	C	2181	0	2166	33	0
1	D	2175	0	2166	29	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	20	0	0	2	1
2	B	5	0	0	0	0
2	C	10	0	0	0	0
2	D	15	0	0	0	0
3	B	18	0	0	0	1
4	A	3	0	0	0	0
4	B	2	0	0	0	0
4	D	1	0	0	0	0
All	All	8740	0	8605	99	2

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 6.

The worst 5 of 99 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:131:VAL:HG13	1:A:136:LEU:HB2	1.66	0.77
1:C:112:PRO:HG3	1:C:118:VAL:HB	1.68	0.73
1:A:114:LYS:HE2	1:B:100:ALA:HA	1.67	0.73
1:C:106:ARG:HA	1:C:109:GLN:HE21	1.54	0.71
1:C:172:LYS:O	1:C:200:ARG:NH2	2.25	0.69

All (2) 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)
2:A:302:PO4:O3	3:B:302:DPO:O7[4_444]	1.26	0.94
1:A:201:ARG:NH2	1:D:183:LEU:O[4_445]	2.18	0.02

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	256/267 (96%)	246 (96%)	8 (3%)	2 (1%)	19	51
1	B	254/267 (95%)	246 (97%)	6 (2%)	2 (1%)	19	51
1	C	254/267 (95%)	245 (96%)	9 (4%)	0	100	100
1	D	254/267 (95%)	246 (97%)	8 (3%)	0	100	100
All	All	1018/1068 (95%)	983 (97%)	31 (3%)	4 (0%)	34	66

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	122	ARG
1	A	122	ARG
1	A	257	PRO
1	B	194	MET

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	227/235 (97%)	223 (98%)	4 (2%)	59	85
1	B	225/235 (96%)	221 (98%)	4 (2%)	59	85
1	C	230/235 (98%)	224 (97%)	6 (3%)	46	77
1	D	229/235 (97%)	225 (98%)	4 (2%)	60	86
All	All	911/940 (97%)	893 (98%)	18 (2%)	55	82

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	15	LYS
1	D	239	VAL
1	D	217	THR
1	C	2	LYS
1	C	200	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such

sidechains are listed below:

Mol	Chain	Res	Type
1	C	109	GLN

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)

12 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	A	303	-	4,4,4	1.20	1 (25%)	6,6,6	0.44	0
2	PO4	C	302	-	4,4,4	1.20	1 (25%)	6,6,6	0.43	0
2	PO4	D	303	-	4,4,4	1.21	1 (25%)	6,6,6	0.44	0
2	PO4	C	301	-	4,4,4	1.17	1 (25%)	6,6,6	0.44	0
2	PO4	A	301	-	4,4,4	1.18	1 (25%)	6,6,6	0.43	0
3	DPO	B	303	-	6,8,8	0.63	0	13,13,13	1.03	2 (15%)
2	PO4	A	304	-	4,4,4	1.18	1 (25%)	6,6,6	0.42	0
2	PO4	A	302	-	4,4,4	1.55	1 (25%)	6,6,6	0.55	0
2	PO4	D	301	-	4,4,4	1.14	1 (25%)	6,6,6	0.42	0
2	PO4	D	302	1	4,4,4	1.12	1 (25%)	6,6,6	0.49	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	DPO	B	302	-	6,8,8	0.70	0	13,13,13	1.07	1 (7%)
2	PO4	B	301	-	4,4,4	1.15	1 (25%)	6,6,6	0.43	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	DPO	B	303	-	-	2/6/6/6	-
3	DPO	B	302	-	-	0/6/6/6	-

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	302	PO4	P-O1	2.77	1.57	1.50
2	D	303	PO4	P-O1	2.25	1.56	1.50
2	A	303	PO4	P-O1	2.24	1.56	1.50
2	C	302	PO4	P-O1	2.24	1.56	1.50
2	C	301	PO4	P-O1	2.22	1.56	1.50

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	302	DPO	O2-P1-O4	2.93	114.48	104.64
3	B	303	DPO	O7-P2-O4	2.53	113.12	104.64
3	B	303	DPO	O2-P1-O4	2.42	112.77	104.64

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	303	DPO	P2-O4-P1-O2
3	B	303	DPO	P2-O4-P1-O1

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	304	PO4	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	302	PO4	0	1
3	B	302	DPO	0	1

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	258/267 (96%)	0.19	1 (0%)	92	93	44, 53, 87, 112	0
1	B	256/267 (95%)	0.21	3 (1%)	79	79	45, 57, 92, 128	0
1	C	260/267 (97%)	0.30	10 (3%)	40	36	49, 68, 110, 134	0
1	D	260/267 (97%)	0.39	18 (6%)	16	13	49, 69, 104, 141	0
All	All	1034/1068 (96%)	0.27	32 (3%)	49	44	44, 62, 100, 141	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	136	LEU	5.0
1	D	4	TYR	3.9
1	B	7	GLN	3.4
1	D	197	LEU	3.3
1	D	135	ASN	3.2

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, 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(Å^2)	Q<0.9
2	PO4	D	302	5/5	0.08	0.54	154,174,175,175	0
2	PO4	C	302	5/5	0.62	0.26	129,131,132,136	0
2	PO4	A	302	5/5	0.74	0.19	109,111,116,123	0
3	DPO	B	302	9/9	0.79	0.38	101,107,124,140	0
3	DPO	B	303	9/9	0.80	0.25	113,128,136,138	0
2	PO4	A	303	5/5	0.85	0.20	112,114,119,119	0
2	PO4	D	303	5/5	0.85	0.16	120,124,126,136	0
2	PO4	A	304	5/5	0.88	0.18	94,96,102,103	0
2	PO4	C	301	5/5	0.93	0.12	58,63,64,68	0
2	PO4	A	301	5/5	0.95	0.18	51,51,52,52	0
2	PO4	D	301	5/5	0.97	0.17	59,60,61,61	0
2	PO4	B	301	5/5	0.98	0.18	55,55,56,57	0

6.5 Other polymers (i)

There are no such residues in this entry.