



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 26, 2017 – 11:57 PM EDT

PDB ID : 1FXP
Title : AQUIFEX AEOLICUS KDO8P SYNTHASE IN COMPLEX WITH CADMIUM
Authors : Duewel, H.S.; Radaev, S.; Wang, J.; Woodard, R.W.; Gatti, D.L.
Deposited on : unknown
Resolution : 1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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-20030345

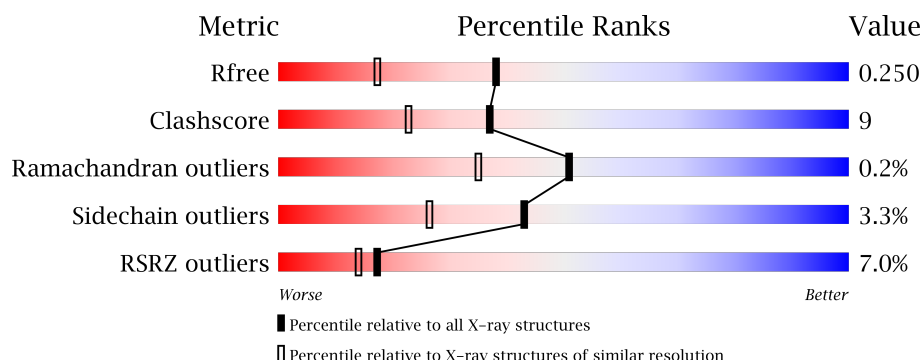
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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}	100719	4827 (1.80-1.80)
Clashscore	112137	5742 (1.80-1.80)
Ramachandran outliers	110173	5676 (1.80-1.80)
Sidechain outliers	110143	5675 (1.80-1.80)
RSRZ outliers	101464	4906 (1.80-1.80)

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	267	<div> <div>3%</div> <div>76%</div> <div>18%</div> <div>5%</div> </div>
1	B	267	<div> <div>10%</div> <div>70%</div> <div>22%</div> <div>6%</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
2	PO4	A	1268	-	-	-	X
2	PO4	B	2268	-	-	-	X

2 Entry composition [i](#)

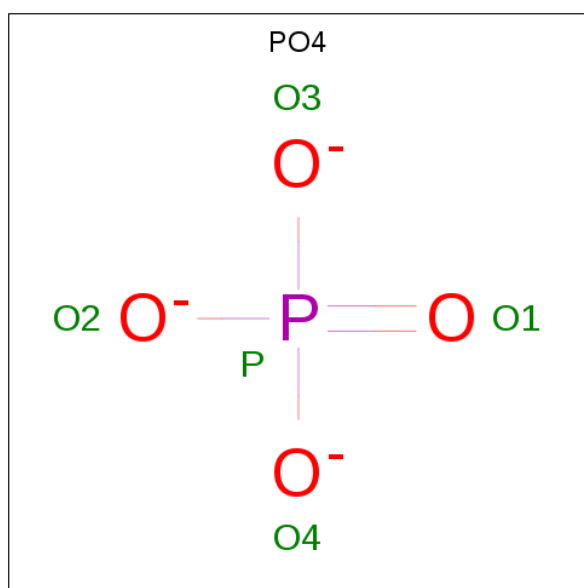
There are 4 unique types of molecules in this entry. The entry contains 4188 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 2-DEHYDRO-3-DEOXYPHOSPHOONATE ALDOLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	253	Total	C	N	O	S	0	0	0
			2006	1297	334	370	5			
1	B	250	Total	C	N	O	S	0	0	0
			1982	1281	331	365	5			

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



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

- Molecule 3 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cd	0	0
			1	1		
3	A	1	Total	Cd	0	0
			1	1		

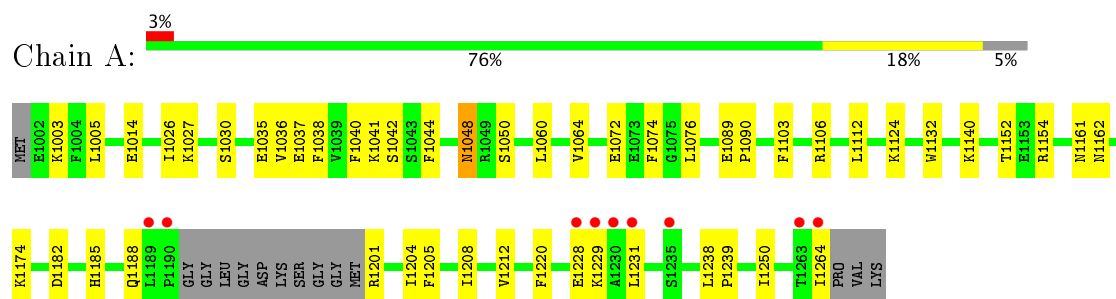
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	101	Total	O	0	0
			101	101		
4	B	77	Total	O	0	0
			77	77		

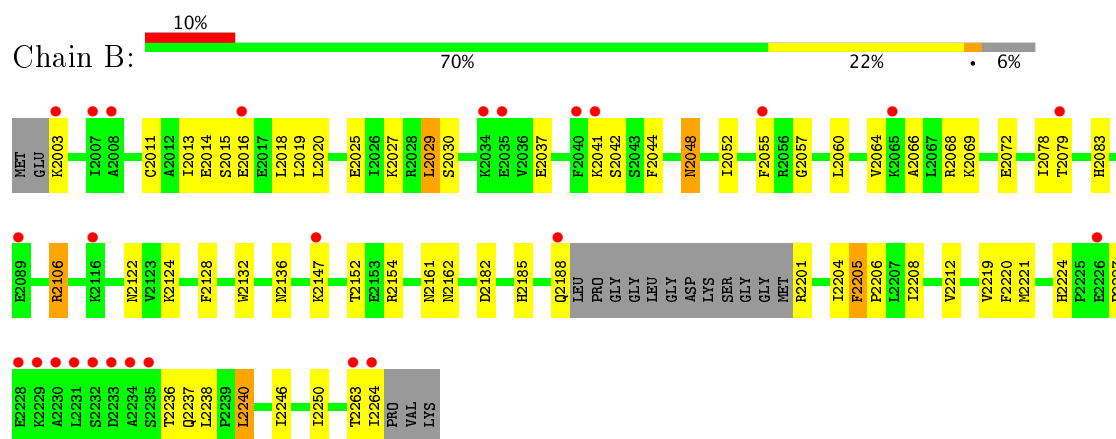
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: 2-DEHYDRO-3-DEOXYPHOSPHOCTONATE ALDOLASE



• Molecule 1: 2-DEHYDRO-3-DEOXYPHOSPHOCTONATE ALDOLASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	84.74Å 84.74Å 159.65Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	26.20 – 1.80 26.20 – 1.80	Depositor EDS
% Data completeness (in resolution range)	83.9 (26.20-1.80) 84.0 (26.20-1.80)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.59 (at 1.80Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.218 , 0.249 0.218 , 0.250	Depositor DCC
R_{free} test set	5327 reflections (10.19%)	DCC
Wilson B-factor (Å ²)	36.8	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 45.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4188	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.63% 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 [i](#)

5.1 Standard geometry [i](#)

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

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	0.37	1/2046 (0.0%)	0.57	0/2761
1	B	0.33	0/2021	0.56	0/2726
All	All	0.35	1/4067 (0.0%)	0.57	0/5487

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1072	GLU	CD-OE2	7.12	1.33	1.25

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	2006	0	2042	34	0
1	B	1982	0	2018	44	0
2	A	10	0	0	0	0
2	B	10	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	101	0	0	2	0
4	B	77	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	4188	0	4060	72	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 9.

All (72) 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:1185:HIS:O	1:A:1188:GLN:HG2	1.73	0.89
1:A:1174:LYS:HG3	1:A:1264:ILE:HD13	1.56	0.88
1:A:1161:ASN:HD21	1:B:2132:TRP:HE1	1.22	0.86
1:B:2208:ILE:HD11	1:B:2219:VAL:HG11	1.68	0.75
1:B:2025:GLU:O	1:B:2029:LEU:HD22	1.86	0.75
1:B:2011:CYS:H	1:B:2237:GLN:NE2	1.88	0.71
1:A:1132:TRP:HE1	1:B:2161:ASN:HD21	1.38	0.70
1:B:2238:LEU:HD21	1:B:2246:ILE:HD12	1.73	0.69
1:A:1201:ARG:HG3	1:A:1204:ILE:HD12	1.73	0.68
1:B:2025:GLU:HG2	1:B:2240:LEU:HD22	1.76	0.67
1:B:2147:LYS:NZ	1:B:2147:LYS:HB3	2.10	0.66
1:B:2205:PHE:O	1:B:2208:ILE:HG22	1.95	0.65
1:B:2185:HIS:O	1:B:2188:GLN:HG2	1.96	0.65
1:B:2208:ILE:CD1	1:B:2219:VAL:HG11	2.27	0.65
4:A:3134:HOH:O	1:B:2083:HIS:HD2	1.84	0.60
1:B:2016:GLU:O	1:B:2020:LEU:HG	2.03	0.59
1:A:1205:PHE:O	1:A:1208:ILE:HG22	2.04	0.57
1:B:2011:CYS:H	1:B:2237:GLN:HE22	1.52	0.55
1:A:1174:LYS:HG3	1:A:1264:ILE:HG21	1.91	0.53
1:A:1048:ASN:HD22	1:A:1048:ASN:H	1.57	0.53
1:B:2205:PHE:HB3	1:B:2206:PRO:HD3	1.90	0.52
1:B:2055:PHE:CZ	1:B:2057:GLY:HA2	2.44	0.52
1:A:1003:LYS:HG3	1:A:1035:GLU:O	2.10	0.52
1:A:1212:VAL:HG11	1:A:1250:ILE:HB	1.91	0.52
1:B:2201:ARG:HG3	1:B:2204:ILE:HD12	1.92	0.51
1:B:2041:LYS:HD3	1:B:2041:LYS:C	2.32	0.50
1:B:2041:LYS:HD3	1:B:2042:SER:N	2.25	0.50
1:A:1124:LYS:HD3	1:A:1152:THR:HB	1.94	0.50
1:A:1014:GLU:HG2	1:A:1231:LEU:HD12	1.93	0.49
1:B:2068:ARG:O	1:B:2072:GLU:HG3	2.13	0.49
1:A:1044:PHE:CZ	1:A:1060:LEU:HD13	2.48	0.49
1:A:1026:ILE:HD12	1:A:1040:PHE:CD1	2.48	0.49
1:A:1060:LEU:HD12	1:A:1064:VAL:HG23	1.95	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1005:LEU:HD12	1:A:1037:GLU:O	2.12	0.48
1:B:2124:LYS:HD3	1:B:2152:THR:HB	1.94	0.48
1:B:2048:ASN:HD22	1:B:2048:ASN:H	1.60	0.47
1:B:2013:ILE:HG23	1:B:2019:LEU:HD11	1.97	0.47
1:B:2147:LYS:HZ2	1:B:2147:LYS:HB3	1.79	0.47
1:B:2003:LYS:HD2	1:B:2037:GLU:HB2	1.96	0.47
1:A:1050:SER:O	1:B:2136:ASN:HB3	2.15	0.47
1:B:2044:PHE:CZ	1:B:2060:LEU:HD13	2.50	0.46
1:A:1003:LYS:HE2	1:A:1036:VAL:C	2.36	0.46
1:A:1041:LYS:C	1:A:1041:LYS:HD3	2.36	0.45
1:A:1027:LYS:HD3	1:A:1076:LEU:HD11	1.99	0.45
1:A:1044:PHE:CE2	1:A:1060:LEU:HD13	2.51	0.45
1:A:1103:PHE:HE2	1:B:2106:ARG:NH1	2.15	0.45
1:B:2027:LYS:O	1:B:2030:SER:HB3	2.17	0.45
1:A:1041:LYS:HD3	1:A:1042:SER:N	2.33	0.44
1:B:2263:THR:HG22	1:B:2264:ILE:HG22	1.99	0.44
1:A:1238:LEU:HD12	1:A:1239:PRO:HD2	2.00	0.44
1:B:2020:LEU:HD21	1:B:2069:LYS:HG3	2.00	0.43
1:A:1074:PHE:HB3	1:A:1076:LEU:HD13	2.00	0.43
1:A:1026:ILE:HD12	1:A:1040:PHE:HD1	1.83	0.43
1:B:2014:GLU:OE1	1:B:2018:LEU:HD22	2.18	0.43
1:B:2060:LEU:HD12	1:B:2064:VAL:HG23	2.00	0.43
1:A:1140:LYS:HG2	1:B:2052:ILE:HB	2.01	0.43
1:A:1089:GLU:HB3	1:A:1090:PRO:CD	2.49	0.43
1:A:1228:GLU:HG3	1:A:1229:LYS:HG3	2.01	0.43
1:B:2019:LEU:HD12	1:B:2066:ALA:HB1	2.00	0.42
1:A:1048:ASN:O	1:B:2106:ARG:NH1	2.52	0.42
1:B:2048:ASN:HD22	1:B:2048:ASN:N	2.17	0.42
1:A:1048:ASN:HD22	1:A:1048:ASN:N	2.15	0.42
4:A:3008:HOH:O	1:B:2128:PHE:HB3	2.20	0.41
1:B:2212:VAL:HG11	1:B:2250:ILE:HB	2.02	0.41
1:B:2188:GLN:HB3	1:B:2236:THR:HG21	2.02	0.41
1:B:2182:ASP:HA	1:B:2220:PHE:HB3	2.02	0.41
1:A:1003:LYS:HE2	1:A:1036:VAL:O	2.21	0.41
1:A:1030:SER:HA	1:A:1038:PHE:CE1	2.56	0.41
1:B:2136:ASN:HA	1:B:2136:ASN:HD22	1.69	0.40
1:B:2224:HIS:O	1:B:2227:PRO:HD3	2.21	0.40
1:A:1182:ASP:HA	1:A:1220:PHE:HB3	2.04	0.40
1:B:2078:ILE:HG22	1:B:2079:THR:N	2.35	0.40

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	249/267 (93%)	245 (98%)	4 (2%)	0	100	100
1	B	246/267 (92%)	241 (98%)	4 (2%)	1 (0%)	38	23
All	All	495/534 (93%)	486 (98%)	8 (2%)	1 (0%)	51	35

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	2015	SER

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	215/224 (96%)	210 (98%)	5 (2%)	56	41
1	B	212/224 (95%)	203 (96%)	9 (4%)	34	17
All	All	427/448 (95%)	413 (97%)	14 (3%)	43	26

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1048	ASN
1	A	1106	ARG
1	A	1112	LEU
1	A	1154	ARG
1	A	1162	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	2029	LEU
1	B	2048	ASN
1	B	2106	ARG
1	B	2122	ASN
1	B	2154	ARG
1	B	2162	ASN
1	B	2205	PHE
1	B	2221	MET
1	B	2240	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1048	ASN
1	A	1058	HIS
1	A	1099	GLN
1	A	1161	ASN
1	A	1162	ASN
1	A	1188	GLN
1	B	2048	ASN
1	B	2083	HIS
1	B	2122	ASN
1	B	2136	ASN
1	B	2161	ASN
1	B	2162	ASN
1	B	2237	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 carbohydrates in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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
2	PO4	A	1268	-	4,4,4	1.34	0	6,6,6	0.38	0
2	PO4	A	1269	-	4,4,4	1.33	0	6,6,6	0.40	0
2	PO4	B	2268	-	4,4,4	1.30	0	6,6,6	0.39	0
2	PO4	B	2269	-	4,4,4	1.33	0	6,6,6	0.39	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
2	PO4	A	1268	-	-	0/0/0/0	0/0/0/0
2	PO4	A	1269	-	-	0/0/0/0	0/0/0/0
2	PO4	B	2268	-	-	0/0/0/0	0/0/0/0
2	PO4	B	2269	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

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.

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	253/267 (94%)	0.08	9 (3%) 43 38	26, 36, 56, 63	0
1	B	250/267 (93%)	0.37	26 (10%) 7 5	28, 41, 62, 78	0
All	All	503/534 (94%)	0.22	35 (6%) 17 14	26, 38, 59, 78	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2264	ILE	5.5
1	A	1264	ILE	5.4
1	A	1235	SER	5.3
1	B	2263	THR	4.8
1	B	2234	ALA	4.7
1	A	1228	GLU	4.2
1	B	2235	SER	4.1
1	B	2231	LEU	3.8
1	B	2147	LYS	3.7
1	B	2055	PHE	3.6
1	A	1189	LEU	3.6
1	B	2228	GLU	3.3
1	A	1263	THR	3.3
1	B	2188	GLN	3.2
1	B	2040	PHE	2.9
1	B	2008	ALA	2.9
1	B	2007	ILE	2.8
1	B	2229	LYS	2.8
1	B	2233	ASP	2.7
1	B	2232	SER	2.6
1	B	2230	ALA	2.5
1	B	2079	THR	2.5
1	A	1231	LEU	2.5
1	B	2226	GLU	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	2089	GLU	2.4
1	A	1229	LYS	2.3
1	B	2116	LYS	2.3
1	B	2065	LYS	2.3
1	B	2034	LYS	2.2
1	A	1230	ALA	2.1
1	B	2035	GLU	2.1
1	B	2003	LYS	2.1
1	B	2041	LYS	2.1
1	A	1190	PRO	2.0
1	B	2016	GLU	2.0

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(Å ²)	Q<0.9
2	PO4	A	1268	5/5	0.87	0.36	9.63	30,30,36,36	5
2	PO4	B	2268	5/5	0.85	0.30	4.02	31,33,38,38	5
2	PO4	B	2269	5/5	0.92	0.14	1.88	49,51,52,53	5
2	PO4	A	1269	5/5	0.91	0.16	1.01	41,42,43,44	5
3	CD	B	2270	1/1	0.96	0.06	-1.18	59,59,59,59	1
3	CD	A	1270	1/1	0.99	0.04	-2.31	52,52,52,52	1

6.5 Other polymers [i](#)

There are no such residues in this entry.