wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 10:12 PM GMT

PDB ID : 4WY0
Title : PdxS (G. stearothermophilus) co-crystallized with R5P in the presence of ammonia.
Authors : Smith, J.L.; Smith, A.M.
Deposited on : 2014-11-14
Resolution : 2.30 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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

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 2.30 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.
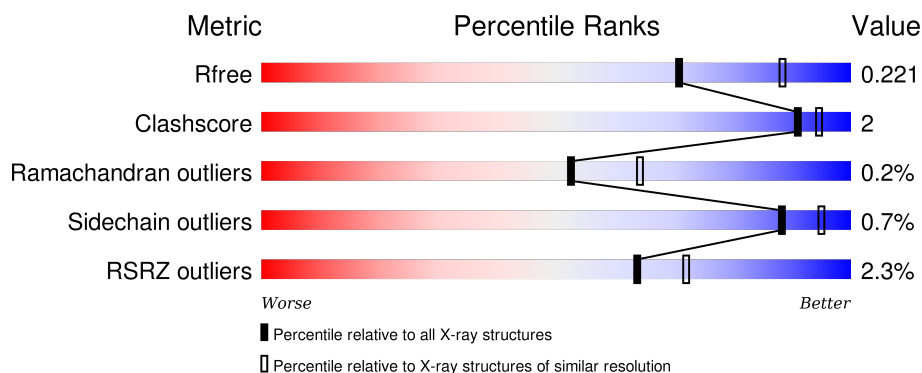
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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 <=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	305	3% 75% 6% 19%
1	C	305	2% 76% 5% 19%
1	E	305	3% 78% • 19%
1	G	305	3% 78% 7% 16%
1	I	305	% 78% • 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	J	305	2% / 78% / 20%
1	K	305	% / 79% / 19%
1	L	305	% / 77% / 19%
2	B	305	2% / 79% / 6% / 15%
2	D	305	% / 79% / 19%
2	H	305	% / 75% / 5% / 19%
3	F	304	2% / 78% / 18%

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
6	R5P	G	301	-	-	-	X
6	R5P	J	301	-	-	-	X
6	R5P	K	301	-	-	-	X
7	EDO	C	302	-	-	-	X
7	EDO	G	302	-	-	X	X
7	EDO	G	303	-	-	-	X

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 23986 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 Pyridoxal biosynthesis lyase PdxS.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	248	Total	C	N	O	P	S	0	0	0
			1885	1177	332	362	1	13			
1	C	247	Total	C	N	O	P	S	0	0	0
			1866	1165	328	359	1	13			
1	E	247	Total	C	N	O	P	S	0	0	0
			1863	1163	328	359	1	12			
1	G	257	Total	C	N	O	P	S	0	0	0
			1955	1218	348	374	1	14			
1	I	249	Total	C	N	O	P	S	0	0	0
			1884	1176	333	361	1	13			
1	J	245	Total	C	N	O	P	S	0	0	0
			1857	1161	326	357	1	12			
1	K	248	Total	C	N	O	P	S	0	0	0
			1885	1177	332	362	1	13			
1	L	246	Total	C	N	O	P	S	0	0	0
			1858	1162	326	357	1	12			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-9	GLU	-	expression tag	UNP Q5L3Y2
A	-8	ASN	-	expression tag	UNP Q5L3Y2
A	-7	LEU	-	expression tag	UNP Q5L3Y2
A	-6	THR	-	expression tag	UNP Q5L3Y2
A	-5	PRO	-	expression tag	UNP Q5L3Y2
A	-4	GLN	-	expression tag	UNP Q5L3Y2
A	-3	HIS	-	expression tag	UNP Q5L3Y2
A	-2	MET	-	expression tag	UNP Q5L3Y2
A	-1	ALA	-	expression tag	UNP Q5L3Y2
A	0	SER	-	expression tag	UNP Q5L3Y2
A	216	THR	ALA	conflict	UNP Q5L3Y2
A	295	ALA	-	expression tag	UNP Q5L3Y2
C	-9	GLU	-	expression tag	UNP Q5L3Y2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-8	ASN	-	expression tag	UNP Q5L3Y2
C	-7	LEU	-	expression tag	UNP Q5L3Y2
C	-6	THR	-	expression tag	UNP Q5L3Y2
C	-5	PRO	-	expression tag	UNP Q5L3Y2
C	-4	GLN	-	expression tag	UNP Q5L3Y2
C	-3	HIS	-	expression tag	UNP Q5L3Y2
C	-2	MET	-	expression tag	UNP Q5L3Y2
C	-1	ALA	-	expression tag	UNP Q5L3Y2
C	0	SER	-	expression tag	UNP Q5L3Y2
C	216	THR	ALA	conflict	UNP Q5L3Y2
C	295	ALA	-	expression tag	UNP Q5L3Y2
E	-9	GLU	-	expression tag	UNP Q5L3Y2
E	-8	ASN	-	expression tag	UNP Q5L3Y2
E	-7	LEU	-	expression tag	UNP Q5L3Y2
E	-6	THR	-	expression tag	UNP Q5L3Y2
E	-5	PRO	-	expression tag	UNP Q5L3Y2
E	-4	GLN	-	expression tag	UNP Q5L3Y2
E	-3	HIS	-	expression tag	UNP Q5L3Y2
E	-2	MET	-	expression tag	UNP Q5L3Y2
E	-1	ALA	-	expression tag	UNP Q5L3Y2
E	0	SER	-	expression tag	UNP Q5L3Y2
E	216	THR	ALA	conflict	UNP Q5L3Y2
E	295	ALA	-	expression tag	UNP Q5L3Y2
G	-9	GLU	-	expression tag	UNP Q5L3Y2
G	-8	ASN	-	expression tag	UNP Q5L3Y2
G	-7	LEU	-	expression tag	UNP Q5L3Y2
G	-6	THR	-	expression tag	UNP Q5L3Y2
G	-5	PRO	-	expression tag	UNP Q5L3Y2
G	-4	GLN	-	expression tag	UNP Q5L3Y2
G	-3	HIS	-	expression tag	UNP Q5L3Y2
G	-2	MET	-	expression tag	UNP Q5L3Y2
G	-1	ALA	-	expression tag	UNP Q5L3Y2
G	0	SER	-	expression tag	UNP Q5L3Y2
G	216	THR	ALA	conflict	UNP Q5L3Y2
G	295	ALA	-	expression tag	UNP Q5L3Y2
I	-9	GLU	-	expression tag	UNP Q5L3Y2
I	-8	ASN	-	expression tag	UNP Q5L3Y2
I	-7	LEU	-	expression tag	UNP Q5L3Y2
I	-6	THR	-	expression tag	UNP Q5L3Y2
I	-5	PRO	-	expression tag	UNP Q5L3Y2
I	-4	GLN	-	expression tag	UNP Q5L3Y2
I	-3	HIS	-	expression tag	UNP Q5L3Y2

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
I	-2	MET	-	expression tag	UNP Q5L3Y2
I	-1	ALA	-	expression tag	UNP Q5L3Y2
I	0	SER	-	expression tag	UNP Q5L3Y2
I	216	THR	ALA	conflict	UNP Q5L3Y2
I	295	ALA	-	expression tag	UNP Q5L3Y2
J	-9	GLU	-	expression tag	UNP Q5L3Y2
J	-8	ASN	-	expression tag	UNP Q5L3Y2
J	-7	LEU	-	expression tag	UNP Q5L3Y2
J	-6	THR	-	expression tag	UNP Q5L3Y2
J	-5	PRO	-	expression tag	UNP Q5L3Y2
J	-4	GLN	-	expression tag	UNP Q5L3Y2
J	-3	HIS	-	expression tag	UNP Q5L3Y2
J	-2	MET	-	expression tag	UNP Q5L3Y2
J	-1	ALA	-	expression tag	UNP Q5L3Y2
J	0	SER	-	expression tag	UNP Q5L3Y2
J	216	THR	ALA	conflict	UNP Q5L3Y2
J	295	ALA	-	expression tag	UNP Q5L3Y2
K	-9	GLU	-	expression tag	UNP Q5L3Y2
K	-8	ASN	-	expression tag	UNP Q5L3Y2
K	-7	LEU	-	expression tag	UNP Q5L3Y2
K	-6	THR	-	expression tag	UNP Q5L3Y2
K	-5	PRO	-	expression tag	UNP Q5L3Y2
K	-4	GLN	-	expression tag	UNP Q5L3Y2
K	-3	HIS	-	expression tag	UNP Q5L3Y2
K	-2	MET	-	expression tag	UNP Q5L3Y2
K	-1	ALA	-	expression tag	UNP Q5L3Y2
K	0	SER	-	expression tag	UNP Q5L3Y2
K	216	THR	ALA	conflict	UNP Q5L3Y2
K	295	ALA	-	expression tag	UNP Q5L3Y2
L	-9	GLU	-	expression tag	UNP Q5L3Y2
L	-8	ASN	-	expression tag	UNP Q5L3Y2
L	-7	LEU	-	expression tag	UNP Q5L3Y2
L	-6	THR	-	expression tag	UNP Q5L3Y2
L	-5	PRO	-	expression tag	UNP Q5L3Y2
L	-4	GLN	-	expression tag	UNP Q5L3Y2
L	-3	HIS	-	expression tag	UNP Q5L3Y2
L	-2	MET	-	expression tag	UNP Q5L3Y2
L	-1	ALA	-	expression tag	UNP Q5L3Y2
L	0	SER	-	expression tag	UNP Q5L3Y2
L	216	THR	ALA	conflict	UNP Q5L3Y2
L	295	ALA	-	expression tag	UNP Q5L3Y2

- Molecule 2 is a protein called Pyridoxal biosynthesis lyase PdxS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	260	Total 1955	C 1220	N 351	O 370	S 14	0	0	0
2	D	247	Total 1857	C 1162	N 331	O 352	S 12	0	0	0
2	H	246	Total 1853	C 1160	N 327	O 353	S 13	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-9	GLU	-	expression tag	UNP Q5L3Y2
B	-8	ASN	-	expression tag	UNP Q5L3Y2
B	-7	LEU	-	expression tag	UNP Q5L3Y2
B	-6	THR	-	expression tag	UNP Q5L3Y2
B	-5	PRO	-	expression tag	UNP Q5L3Y2
B	-4	GLN	-	expression tag	UNP Q5L3Y2
B	-3	HIS	-	expression tag	UNP Q5L3Y2
B	-2	MET	-	expression tag	UNP Q5L3Y2
B	-1	ALA	-	expression tag	UNP Q5L3Y2
B	0	SER	-	expression tag	UNP Q5L3Y2
B	216	THR	ALA	conflict	UNP Q5L3Y2
B	295	ALA	-	expression tag	UNP Q5L3Y2
D	-9	GLU	-	expression tag	UNP Q5L3Y2
D	-8	ASN	-	expression tag	UNP Q5L3Y2
D	-7	LEU	-	expression tag	UNP Q5L3Y2
D	-6	THR	-	expression tag	UNP Q5L3Y2
D	-5	PRO	-	expression tag	UNP Q5L3Y2
D	-4	GLN	-	expression tag	UNP Q5L3Y2
D	-3	HIS	-	expression tag	UNP Q5L3Y2
D	-2	MET	-	expression tag	UNP Q5L3Y2
D	-1	ALA	-	expression tag	UNP Q5L3Y2
D	0	SER	-	expression tag	UNP Q5L3Y2
D	216	THR	ALA	conflict	UNP Q5L3Y2
D	295	ALA	-	expression tag	UNP Q5L3Y2
H	-9	GLU	-	expression tag	UNP Q5L3Y2
H	-8	ASN	-	expression tag	UNP Q5L3Y2
H	-7	LEU	-	expression tag	UNP Q5L3Y2
H	-6	THR	-	expression tag	UNP Q5L3Y2
H	-5	PRO	-	expression tag	UNP Q5L3Y2
H	-4	GLN	-	expression tag	UNP Q5L3Y2
H	-3	HIS	-	expression tag	UNP Q5L3Y2
H	-2	MET	-	expression tag	UNP Q5L3Y2
H	-1	ALA	-	expression tag	UNP Q5L3Y2

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	0	SER	-	expression tag	UNP Q5L3Y2
H	216	THR	ALA	conflict	UNP Q5L3Y2
H	295	ALA	-	expression tag	UNP Q5L3Y2

- Molecule 3 is a protein called Pyridoxal biosynthesis lyase PdxS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	F	248	1877	1177	333	354	13	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-9	GLU	-	expression tag	UNP Q5L3Y2
F	-8	ASN	-	expression tag	UNP Q5L3Y2
F	-7	LEU	-	expression tag	UNP Q5L3Y2
F	-6	THR	-	expression tag	UNP Q5L3Y2
F	-5	PRO	-	expression tag	UNP Q5L3Y2
F	-4	GLN	-	expression tag	UNP Q5L3Y2
F	-3	HIS	-	expression tag	UNP Q5L3Y2
F	-2	MET	-	expression tag	UNP Q5L3Y2
F	-1	ALA	-	expression tag	UNP Q5L3Y2
F	0	SER	-	expression tag	UNP Q5L3Y2
F	216	THR	ALA	conflict	UNP Q5L3Y2

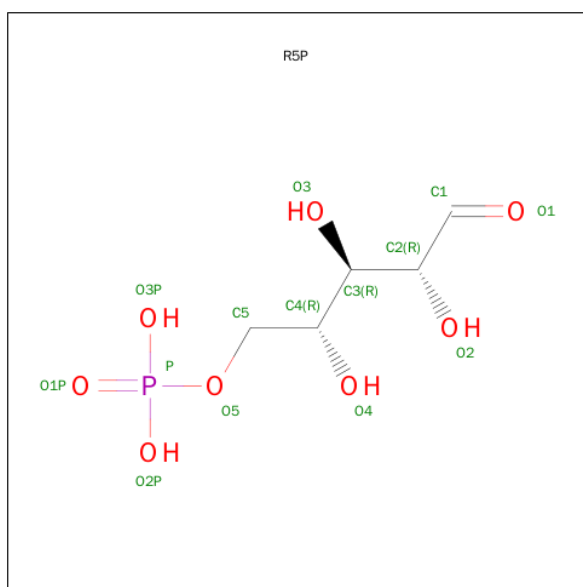- Molecule 4 is CACODYLATE ION (three-letter code: CAC) (formula: $C_2H_6AsO_2$).
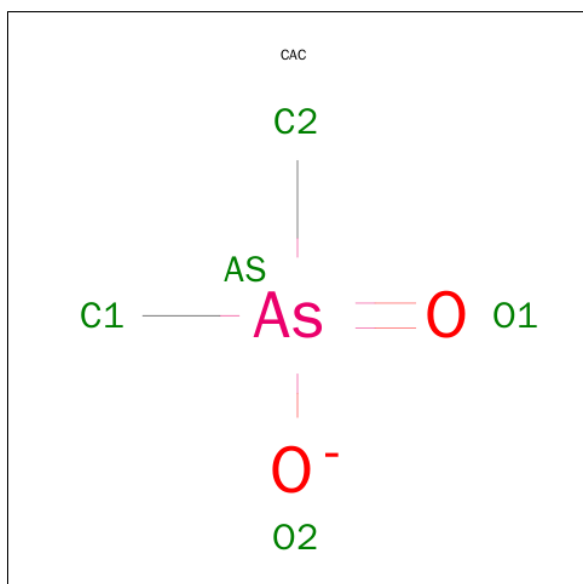
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 5	As 1	C 2	O 2	0	0
4	C	1	Total 5	As 1	C 2	O 2	0	0
4	F	1	Total 5	As 1	C 2	O 2	0	0
4	H	1	Total 5	As 1	C 2	O 2	0	0
4	L	1	Total 5	As 1	C 2	O 2	0	0

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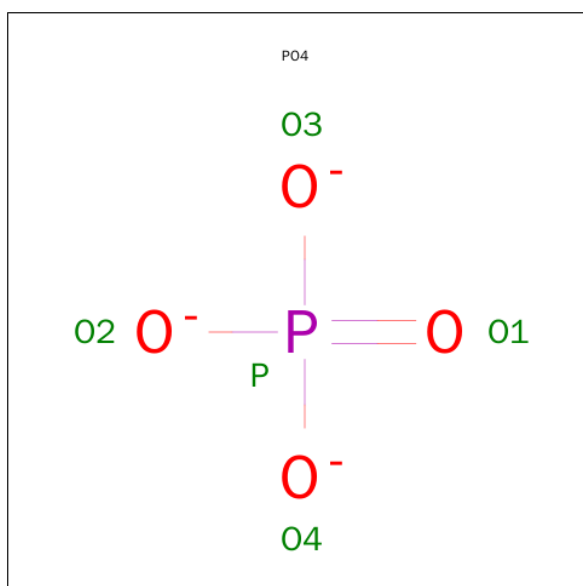
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total 5	O 4	P 1	0	0
5	D	1	Total 5	O 4	P 1	0	0
5	F	1	Total 5	O 4	P 1	0	0
5	H	1	Total 5	O 4	P 1	0	0

- Molecule 6 is SUGAR (RIBOSE-5-PHOSPHATE) (three-letter code: R5P) (formula: $C_5H_{11}O_8P$).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	1	Total 14	C 5	O 8	P 1	0	0
6	D	1	Total 14	C 5	O 8	P 1	0	0
6	E	1	Total 14	C 5	O 8	P 1	0	0
6	G	1	Total 14	C 5	O 8	P 1	0	0
6	I	1	Total 14	C 5	O 8	P 1	0	0
6	J	1	Total 14	C 5	O 8	P 1	0	0
6	K	1	Total 14	C 5	O 8	P 1	0	0

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

EDO
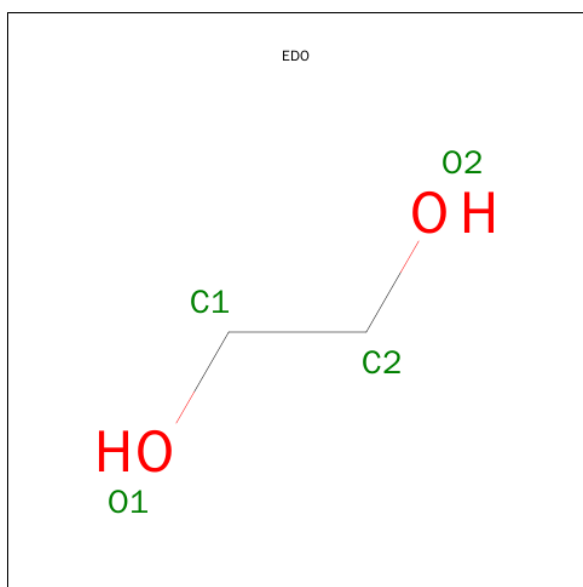
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	C	1	Total 4	C 2	O 2	0	0
7	G	1	Total 4	C 2	O 2	0	0
7	G	1	Total 4	C 2	O 2	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	115	Total 115	O 115	0	0
8	B	114	Total 114	O 114	0	0
8	C	96	Total 96	O 96	0	0
8	D	119	Total 119	O 119	0	0
8	E	102	Total 102	O 102	0	0
8	F	69	Total 69	O 69	0	0
8	G	123	Total 123	O 123	0	0
8	H	109	Total 109	O 109	0	0
8	I	106	Total 106	O 106	0	0

Continued on next page...

Continued from previous page...

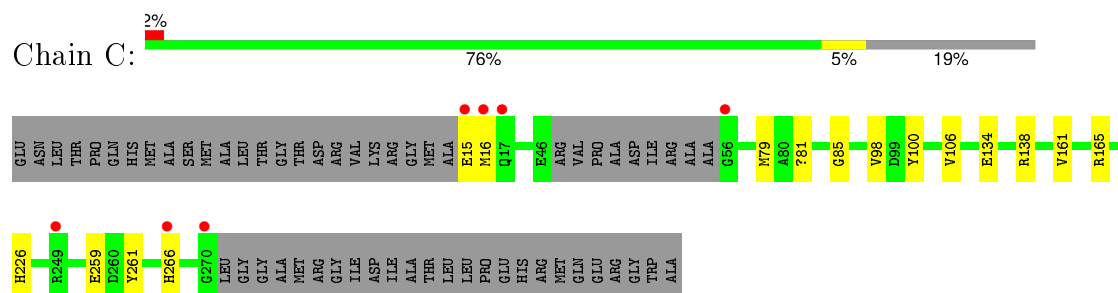| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	J	94	Total 94	O 94	0	0
8	K	104	Total 104	O 104	0	0
8	L	85	Total 85	O 85	0	0

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.
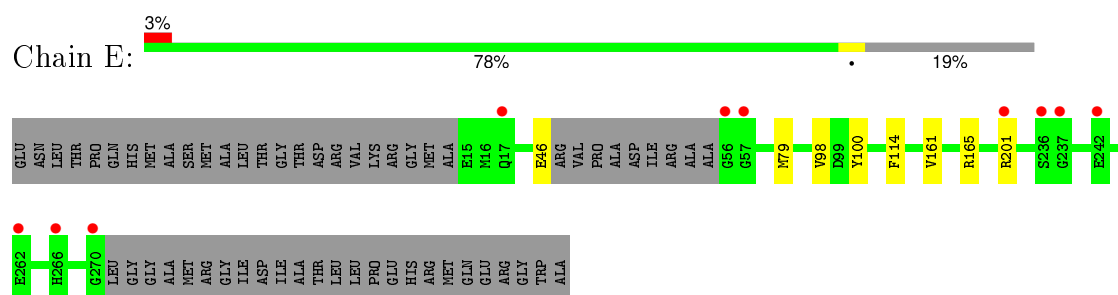
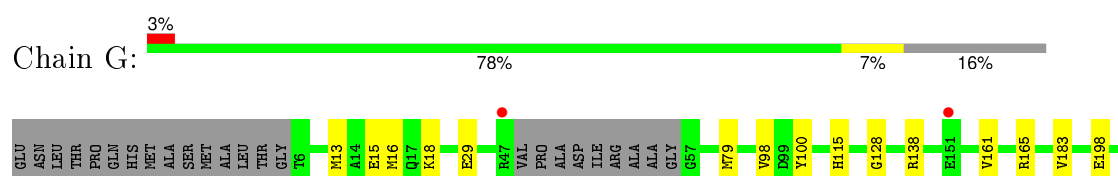- Molecule 1: Pyridoxal biosynthesis lyase PdxS
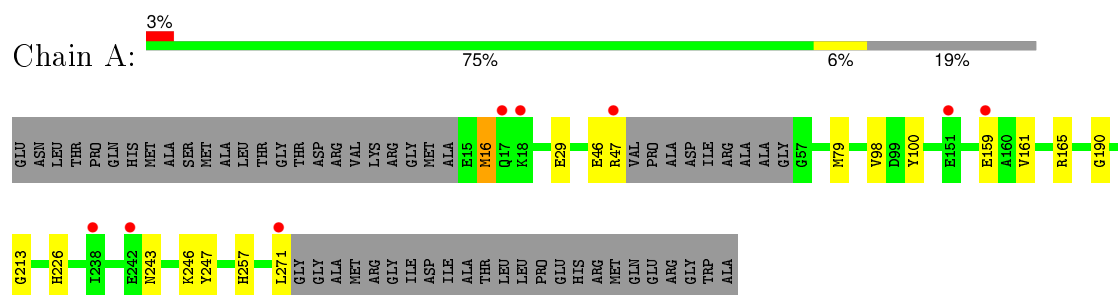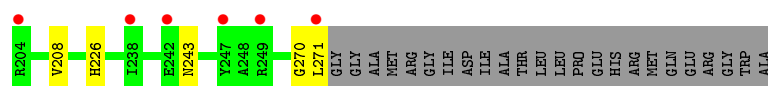

- Molecule 1: Pyridoxal biosynthesis lyase PdxS

- Molecule 1: Pyridoxal biosynthesis lyase PdxS

- Molecule 1: Pyridoxal biosynthesis lyase PdxS

• Molecule 1: Pyridoxal biosynthesis lyase PdxS

Chain I:

• Molecule 1: Pyridoxal biosynthesis lyase PdxS

Chain J:

• Molecule 1: Pyridoxal biosynthesis lyase PdxS

Chain K:

• Molecule 1: Pyridoxal biosynthesis lyase PdxS

Chain L:

• Molecule 2: Pyridoxal biosynthesis lyase PdxS

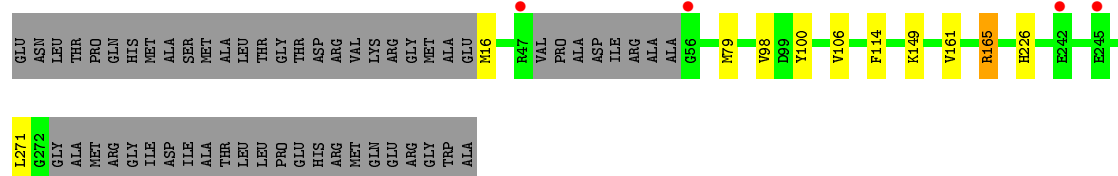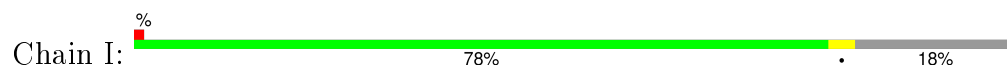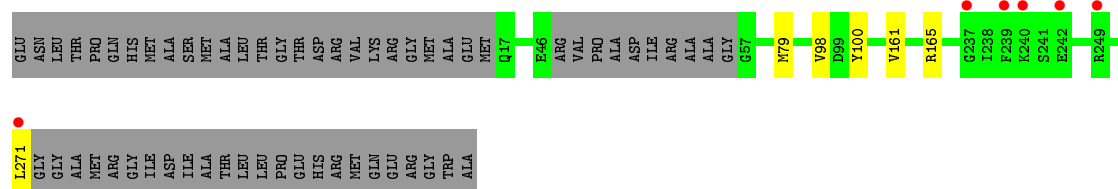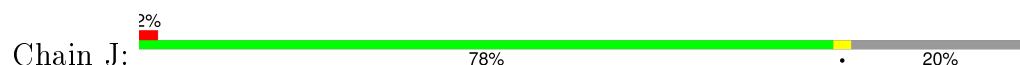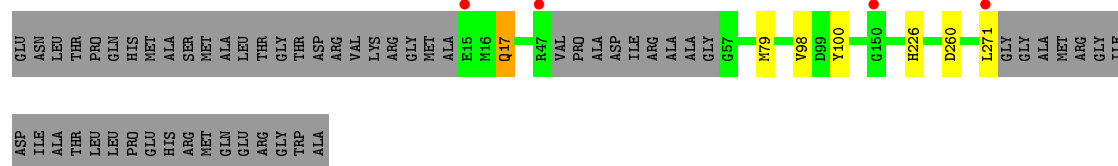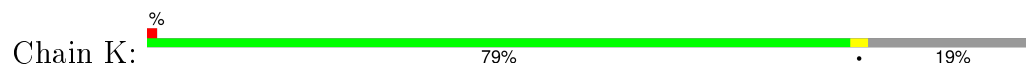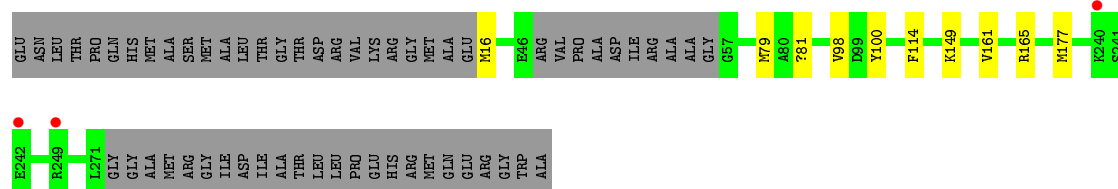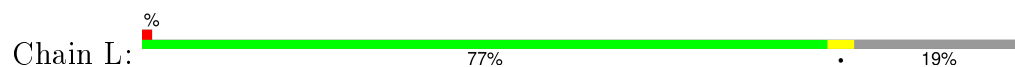Chain B:

• Molecule 2: Pyridoxal biosynthesis lyase PdxS

Chain D:

• Molecule 2: Pyridoxal biosynthesis lyase PdxS

Chain H:
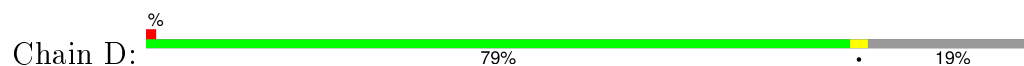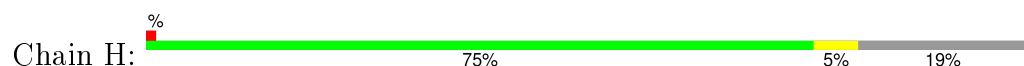
• Molecule 3: Pyridoxal biosynthesis lyase PdxS

Chain F:
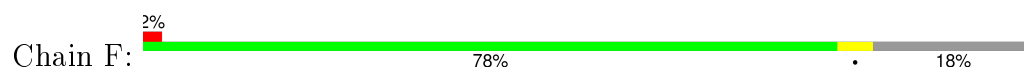
4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	94.68Å 107.20Å 178.22Å 90.00° 92.24° 90.00°	Depositor
Resolution (Å)	50.00 – 2.30 47.30 – 2.29	Depositor EDS
% Data completeness (in resolution range)	99.4 (50.00-2.30) 99.5 (47.30-2.29)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.26 (at 2.29Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.191 , 0.222 0.193 , 0.221	Depositor DCC
R_{free} test set	7934 reflections (5.29%)	DCC
Wilson B-factor (Å^2)	43.2	Xtriage
Anisotropy	0.242	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 34.9	EDS
Estimated twinning fraction	0.014 for h,-k,-l	Xtriage
L-test for twinning[2]	$< |L| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Outliers	0 of 158811 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	23986	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.62% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: L5P, 3ZL, PO4, EDO, CAC, R5P

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	0/1886	0.58	0/2545
1	C	0.36	0/1867	0.58	0/2520
1	E	0.35	0/1864	0.58	0/2517
1	G	0.37	0/1956	0.59	0/2637
1	I	0.36	0/1885	0.60	2/2543 (0.1%)
1	J	0.35	0/1858	0.58	0/2509
1	K	0.35	0/1886	0.58	0/2545
1	L	0.32	0/1859	0.57	0/2511
2	B	0.37	0/1979	0.61	0/2668
2	D	0.36	0/1881	0.58	0/2540
2	H	0.34	0/1877	0.57	0/2534
3	F	0.35	0/1884	0.58	1/2543 (0.0%)
All	All	0.35	0/22682	0.58	3/30612 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	I	165	ARG	NE-CZ-NH1	5.38	122.99	120.30
1	I	165	ARG	NE-CZ-NH2	-5.37	117.62	120.30
3	F	165	ARG	NE-CZ-NH1	5.12	122.86	120.30

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	1885	0	1901	11	2
1	C	1866	0	1876	10	2
1	E	1863	0	1869	6	2
1	G	1955	0	1977	20	2
1	I	1884	0	1901	7	0
1	J	1857	0	1873	2	0
1	K	1885	0	1901	4	1
1	L	1858	0	1869	8	0
2	B	1955	0	1984	15	0
2	D	1857	0	1875	4	0
2	H	1853	0	1873	13	1
3	F	1877	0	1903	8	2
4	A	5	0	0	0	0
4	C	5	0	0	2	0
4	F	5	0	0	0	0
4	H	5	0	0	2	0
4	L	5	0	0	0	0
5	B	5	0	0	0	0
5	D	5	0	0	0	0
5	F	5	0	0	0	0
5	H	5	0	0	0	0
6	B	14	0	9	0	0
6	D	14	0	9	1	0
6	E	14	0	9	1	0
6	G	14	0	9	0	0
6	I	14	0	9	0	0
6	J	14	0	9	0	0
6	K	14	0	9	0	0
7	C	4	0	6	1	0
7	G	8	0	12	10	0
8	A	115	0	0	7	0
8	B	114	0	0	6	0
8	C	96	0	0	3	0
8	D	119	0	0	1	0
8	E	102	0	0	1	0
8	F	69	0	0	1	0
8	G	123	0	0	4	0
8	H	109	0	0	6	0
8	I	106	0	0	2	0
8	J	94	0	0	0	0
8	K	104	0	0	3	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	L	85	0	0	2	0
All	All	23986	0	22883	100	6

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 2.

The worst 5 of 100 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:159:GLU:HB2	8:A:414:HOH:O	1.44	1.17
1:G:128:GLY:O	7:G:302:EDO:H22	1.58	1.02
1:K:260:ASP:HB2	8:K:404:HOH:O	1.85	0.77
2:H:19:GLY:HA3	8:H:401:HOH:O	1.88	0.74
1:G:183:VAL:O	7:G:303:EDO:O2	2.05	0.74

The worst 5 of 6 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)
1:A:243:ASN:OD1	1:G:29:GLU:OE2[1_455]	1.55	0.65
1:A:29:GLU:OE2	1:G:243:ASN:OD1[1_455]	1.62	0.58
1:C:266:HIS:NE2	1:K:260:ASP:OD2[2_354]	1.75	0.45
1:E:201:ARG:CG	3:F:260:ASP:OD2[1_565]	1.88	0.32
1:C:259:GLU:OE1	2:H:32:LYS:CD[2_454]	2.13	0.07

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	242/305 (79%)	234 (97%)	7 (3%)	1 (0%)	39	48
1	C	241/305 (79%)	234 (97%)	7 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	241/305 (79%)	235 (98%)	6 (2%)	0	100	100
1	G	251/305 (82%)	245 (98%)	6 (2%)	0	100	100
1	I	243/305 (80%)	237 (98%)	5 (2%)	1 (0%)	39	48
1	J	239/305 (78%)	234 (98%)	5 (2%)	0	100	100
1	K	242/305 (79%)	234 (97%)	7 (3%)	1 (0%)	39	48
1	L	240/305 (79%)	236 (98%)	4 (2%)	0	100	100
2	B	256/305 (84%)	250 (98%)	5 (2%)	1 (0%)	39	48
2	D	243/305 (80%)	236 (97%)	6 (2%)	1 (0%)	39	48
2	H	242/305 (79%)	235 (97%)	5 (2%)	2 (1%)	24	27
3	F	243/304 (80%)	236 (97%)	7 (3%)	0	100	100
All	All	2923/3659 (80%)	2846 (97%)	70 (2%)	7 (0%)	52	64

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	16	MET
2	H	17	GLN
1	I	271	LEU
1	K	17	GLN
2	H	18	LYS

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	192/233 (82%)	190 (99%)	2 (1%)	82	91
1	C	189/233 (81%)	188 (100%)	1 (0%)	92	97
1	E	188/233 (81%)	187 (100%)	1 (0%)	92	97
1	G	199/233 (85%)	198 (100%)	1 (0%)	92	97
1	I	191/233 (82%)	190 (100%)	1 (0%)	92	97
1	J	189/233 (81%)	187 (99%)	2 (1%)	80	90

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles		
1	K	192/233 (82%)	189 (98%)	3 (2%)		70	84
1	L	188/233 (81%)	187 (100%)	1 (0%)		92	97
2	B	200/234 (86%)	199 (100%)	1 (0%)		92	97
2	D	190/234 (81%)	189 (100%)	1 (0%)		92	97
2	H	191/234 (82%)	190 (100%)	1 (0%)		92	97
3	F	192/233 (82%)	191 (100%)	1 (0%)		92	97
All	All	2301/2799 (82%)	2285 (99%)	16 (1%)		88	95

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	98	VAL
2	H	98	VAL
1	K	17	GLN
3	F	98	VAL
1	K	98	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
3	F	17	GLN
1	G	115	HIS
1	G	226	HIS

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

9 non-standard protein/DNA/RNA residues 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
1	L5P	A	81	1	19,21,22	1.05	1 (5%)	15,27,29	0.82	0
1	L5P	C	81	1	19,21,22	1.13	2 (10%)	15,27,29	0.88	0
1	L5P	E	81	1	19,21,22	0.91	1 (5%)	15,27,29	1.05	1 (6%)
3	3ZL	F	81	3	13,15,16	0.97	1 (7%)	11,17,19	2.51	5 (45%)
1	L5P	G	81	1	19,21,22	1.10	2 (10%)	15,27,29	0.77	0
1	L5P	I	81	1	19,21,22	1.07	1 (5%)	15,27,29	0.85	0
1	L5P	J	81	1	19,21,22	1.14	1 (5%)	15,27,29	0.82	0
1	L5P	K	81	1	19,21,22	1.17	2 (10%)	15,27,29	1.30	2 (13%)
1	L5P	L	81	1	19,21,22	1.12	2 (10%)	15,27,29	0.93	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
1	L5P	A	81	1	-	0/22/25/27	0/0/0/0
1	L5P	C	81	1	-	0/22/25/27	0/0/0/0
1	L5P	E	81	1	-	0/22/25/27	0/0/0/0
3	3ZL	F	81	3	-	0/9/16/18	0/0/0/0
1	L5P	G	81	1	-	0/22/25/27	0/0/0/0
1	L5P	I	81	1	-	0/22/25/27	0/0/0/0
1	L5P	J	81	1	-	0/22/25/27	0/0/0/0
1	L5P	K	81	1	-	0/22/25/27	0/0/0/0
1	L5P	L	81	1	-	0/22/25/27	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	81	L5P	P9-O10	2.03	1.62	1.54
1	E	81	L5P	P9-O10	2.08	1.62	1.54
1	K	81	L5P	P9-O10	2.09	1.62	1.54
1	L	81	L5P	P9-O10	2.10	1.62	1.54
1	C	81	L5P	P9-O10	2.18	1.62	1.54

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	81	3ZL	C5-C4-N14	-3.72	118.03	126.27
1	K	81	L5P	O13-C5-C3	-3.20	105.09	111.01

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	K	81	L5P	O-C1-CA	-2.03	120.20	125.49
3	F	81	3ZL	O-C-CA	-2.01	120.26	125.49
1	E	81	L5P	C7-C6-C5	2.72	116.46	111.85

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	81	L5P	1	0
1	L	81	L5P	2	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

19 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 Counts	Bond lengths RMSZ	Bond lengths #$|Z| > 2$	Bond angles Counts	Bond angles RMSZ	Bond angles #$|Z| > 2$
4	CAC	A	301	-	0,4,4	0.00	-	0,6,6	0.00	-
5	PO4	B	301	-	4,4,4	0.19	0	6,6,6	0.29	0
6	R5P	B	302	-	13,13,13	1.72	2 (15%)	15,18,18	1.61	2 (13%)
4	CAC	C	301	-	0,4,4	0.00	-	0,6,6	0.00	-
7	EDO	C	302	-	3,3,3	0.53	0	2,2,2	0.37	0
5	PO4	D	301	-	4,4,4	0.33	0	6,6,6	0.27	0
6	R5P	D	302	-	13,13,13	1.66	2 (15%)	15,18,18	1.69	3 (20%)
6	R5P	E	301	-	13,13,13	1.60	1 (7%)	15,18,18	1.70	1 (6%)
5	PO4	F	301	-	4,4,4	0.44	0	6,6,6	0.28	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
4	CAC	F	302	-	0,4,4	0.00	-	0,6,6	0.00	-
6	R5P	G	301	-	13,13,13	1.65	1 (7%)	15,18,18	1.50	1 (6%)
7	EDO	G	302	-	3,3,3	0.34	0	2,2,2	0.08	0
7	EDO	G	303	-	3,3,3	0.45	0	2,2,2	0.12	0
5	PO4	H	301	-	4,4,4	0.32	0	6,6,6	0.28	0
4	CAC	H	302	-	0,4,4	0.00	-	0,6,6	0.00	-
6	R5P	I	301	-	13,13,13	1.64	1 (7%)	15,18,18	1.54	1 (6%)
6	R5P	J	301	-	13,13,13	1.67	1 (7%)	15,18,18	1.63	1 (6%)
6	R5P	K	301	-	13,13,13	1.71	2 (15%)	15,18,18	1.68	1 (6%)
4	CAC	L	301	-	0,4,4	0.00	-	0,6,6	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	CAC	A	301	-	-	0/0/0/0	0/0/0/0
5	PO4	B	301	-	-	0/0/0/0	0/0/0/0
6	R5P	B	302	-	-	0/14/16/16	0/0/0/0
4	CAC	C	301	-	-	0/0/0/0	0/0/0/0
7	EDO	C	302	-	-	0/1/1/1	0/0/0/0
5	PO4	D	301	-	-	0/0/0/0	0/0/0/0
6	R5P	D	302	-	-	0/14/16/16	0/0/0/0
6	R5P	E	301	-	-	0/14/16/16	0/0/0/0
5	PO4	F	301	-	-	0/0/0/0	0/0/0/0
4	CAC	F	302	-	-	0/0/0/0	0/0/0/0
6	R5P	G	301	-	-	0/14/16/16	0/0/0/0
7	EDO	G	302	-	-	0/1/1/1	0/0/0/0
7	EDO	G	303	-	-	0/1/1/1	0/0/0/0
5	PO4	H	301	-	-	0/0/0/0	0/0/0/0
4	CAC	H	302	-	-	0/0/0/0	0/0/0/0
6	R5P	I	301	-	-	0/14/16/16	0/0/0/0
6	R5P	J	301	-	-	0/14/16/16	0/0/0/0
6	R5P	K	301	-	-	0/14/16/16	0/0/0/0
4	CAC	L	301	-	-	0/0/0/0	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	K	301	R5P	C2-C1	2.00	1.53	1.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	302	R5P	C2-C1	2.09	1.54	1.50
6	B	302	R5P	C2-C1	2.29	1.54	1.50
6	D	302	R5P	O1-C1	5.20	1.43	1.19
6	G	301	R5P	O1-C1	5.21	1.43	1.19

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	E	301	R5P	O1-C1-C2	-5.26	110.26	125.60
6	K	301	R5P	O1-C1-C2	-5.20	110.45	125.60
6	D	302	R5P	O1-C1-C2	-5.06	110.84	125.60
6	J	301	R5P	O1-C1-C2	-5.02	110.97	125.60
6	I	301	R5P	O1-C1-C2	-4.80	111.61	125.60

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	301	CAC	2	0
7	C	302	EDO	1	0
6	D	302	R5P	1	0
6	E	301	R5P	1	0
7	G	302	EDO	7	0
7	G	303	EDO	3	0
4	H	302	CAC	2	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	247/305 (80%)	-0.11	8 (3%)	51	60	31, 42, 74, 114	0
1	C	246/305 (80%)	-0.11	7 (2%)	56	66	32, 43, 71, 116	0
1	E	246/305 (80%)	-0.04	10 (4%)	41	50	34, 47, 81, 111	0
1	G	256/305 (83%)	-0.12	8 (3%)	52	62	32, 42, 69, 102	0
1	I	248/305 (81%)	-0.07	4 (1%)	74	80	32, 43, 72, 105	0
1	J	244/305 (80%)	-0.11	6 (2%)	61	70	33, 46, 73, 106	0
1	K	247/305 (80%)	-0.09	4 (1%)	74	80	33, 45, 73, 111	0
1	L	245/305 (80%)	-0.06	3 (1%)	81	85	37, 49, 78, 107	0
2	B	260/305 (85%)	0.01	7 (2%)	58	67	32, 43, 80, 108	0
2	D	247/305 (80%)	-0.12	4 (1%)	74	80	34, 43, 74, 115	0
2	H	246/305 (80%)	-0.13	2 (0%)	87	90	35, 48, 79, 96	0
3	F	247/304 (81%)	0.02	7 (2%)	56	66	34, 47, 80, 114	0
All	All	2979/3659 (81%)	-0.08	70 (2%)	64	72	31, 45, 77, 116	0

The worst 5 of 70 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	56	GLY	6.0
1	A	271	LEU	5.6
1	K	271	LEU	4.9
1	E	56	GLY	4.8
1	E	242	GLU	4.1

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. 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
1	L5P	E	81	22/23	0.89	0.16	-	37,87,94,99	0
3	3ZL	F	81	16/17	0.94	0.15	-	41,58,78,79	0
1	L5P	C	81	22/23	0.93	0.13	-	39,73,81,83	0
1	L5P	A	81	22/23	0.95	0.13	-	34,81,87,90	0
1	L5P	L	81	22/23	0.92	0.15	-	44,83,92,95	0
1	L5P	J	81	22/23	0.91	0.13	-	36,77,85,88	0
1	L5P	K	81	22/23	0.92	0.12	-	44,78,85,87	0
1	L5P	I	81	22/23	0.90	0.13	-	39,76,85,87	0
1	L5P	G	81	22/23	0.94	0.12	-	39,72,76,79	0

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
7	EDO	G	303	4/4	0.48	1.15	53.64	89,93,95,99	0
7	EDO	G	302	4/4	0.68	0.72	33.67	67,69,69,70	0
7	EDO	C	302	4/4	0.93	0.33	10.34	56,59,61,61	0
6	R5P	G	301	14/14	0.93	0.23	2.97	63,79,86,86	0
6	R5P	K	301	14/14	0.91	0.20	2.91	60,79,96,96	0
6	R5P	J	301	14/14	0.93	0.19	2.30	56,75,84,85	0
6	R5P	B	302	14/14	0.92	0.19	1.96	57,79,89,90	0
6	R5P	D	302	14/14	0.93	0.18	1.77	55,79,91,92	0
5	PO4	B	301	5/5	0.82	0.17	1.64	61,62,73,75	0
6	R5P	I	301	14/14	0.94	0.19	1.34	58,79,101,102	0
6	R5P	E	301	14/14	0.96	0.17	1.14	59,75,80,83	0
5	PO4	D	301	5/5	0.76	0.17	0.85	76,77,87,93	0
4	CAC	H	302	5/5	0.91	0.14	0.11	79,80,86,89	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
4	CAC	C	301	5/5	0.96	0.12	0.03	80,81,85,91	0
5	PO4	H	301	5/5	0.93	0.11	-0.14	77,81,84,85	0
5	PO4	F	301	5/5	0.94	0.11	-0.26	79,80,85,85	0
4	CAC	A	301	5/5	0.95	0.12	-0.70	79,86,88,88	0
4	CAC	F	302	5/5	0.96	0.10	-1.01	67,77,79,80	0
4	CAC	L	301	5/5	0.96	0.09	-1.63	75,76,81,82	0

6.5 Other polymers ⓘ

There are no such residues in this entry.