W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Sep 11, 2023 – 07:51 PM EDT

PDB ID	:	4LNJ
Title	:	Structure of Escherichia coli Threonine Aldolase in Unliganded Form
Authors	:	Safo, M.K.; Contestabile, R.; Remesh, S.G.
Deposited on	:	2013-07-11
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
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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
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)	:	2.35.1

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	Percentile Ranks	Value
Rfree		0.282
Clashscore		19
Ramachandran outliers		1.2%
Sidechain outliers		3.5%
RSRZ outliers		2.7%

Worse *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
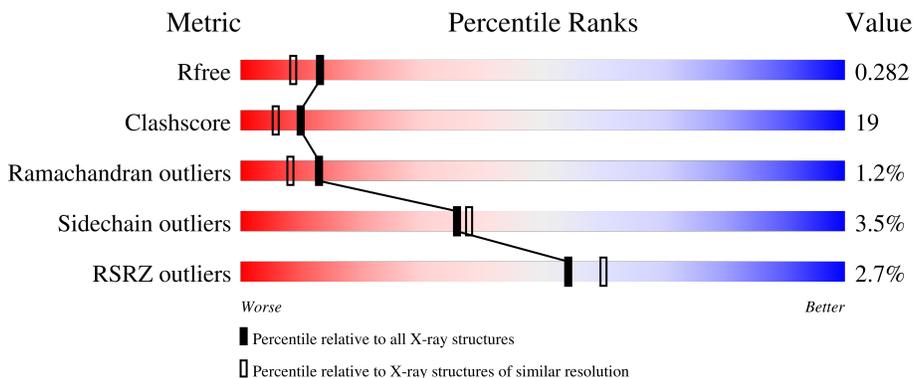
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 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	333	2% ▮ 66% 31% •
1	B	333	4% ▮ 65% 32% • •

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
3	PLR	A	403	-	X	-	-

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 5449 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 Low-specificity L-threonine aldolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	332	Total	C	N	O	S	0	1	0
			2553	1601	460	476	16			
1	B	332	Total	C	N	O	S	0	1	0
			2554	1602	460	476	16			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	256	THR	ALA	conflict	UNP E7U392
A	257	ALA	THR	conflict	UNP E7U392
B	256	THR	ALA	conflict	UNP E7U392
B	257	ALA	THR	conflict	UNP E7U392

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Mg	0	0
			2	2		
2	B	1	Total	Mg	0	0
			1	1		

- Molecule 3 is (5-HYDROXY-4,6-DIMETHYLPYRIDIN-3-YL)METHYL DIHYDROGEN PHOSPHATE (three-letter code: PLR) (formula: $C_8H_{12}NO_5P$).
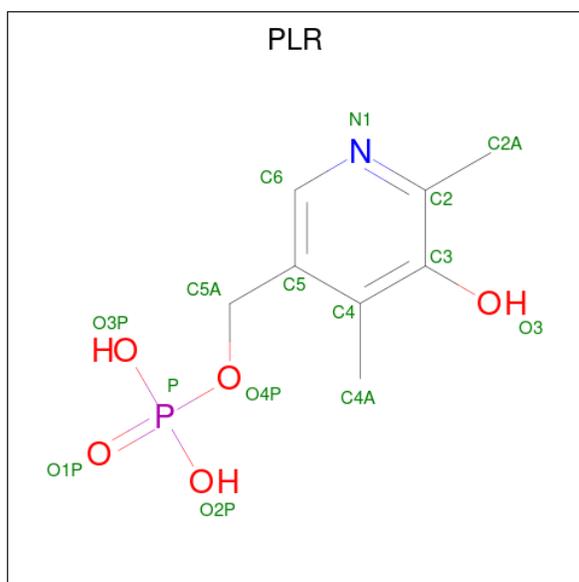
PLR

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
3	B	1	Total	C	N	O	P	0	0
			15	8	1	5	1		

- Molecule 4 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).
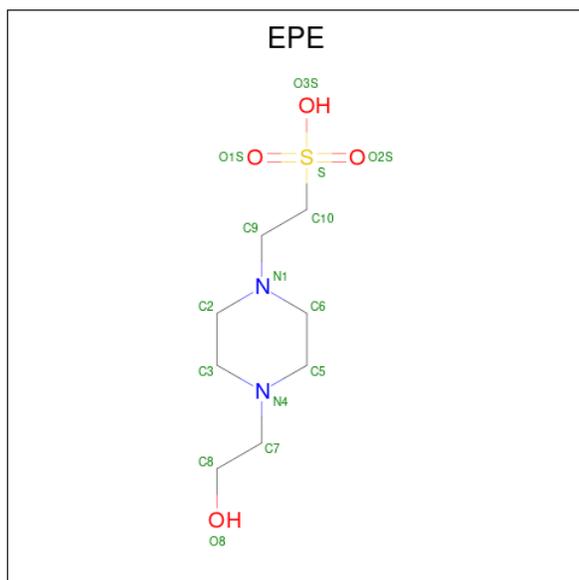
EPE

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total 15	C 8	N 2	O 4	S 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	144	Total 144	O 144	0	0
5	B	135	Total 135	O 135	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Low-specificity L-threonine aldolase

Chain A:

- Molecule 1: Low-specificity L-threonine aldolase

Chain B:
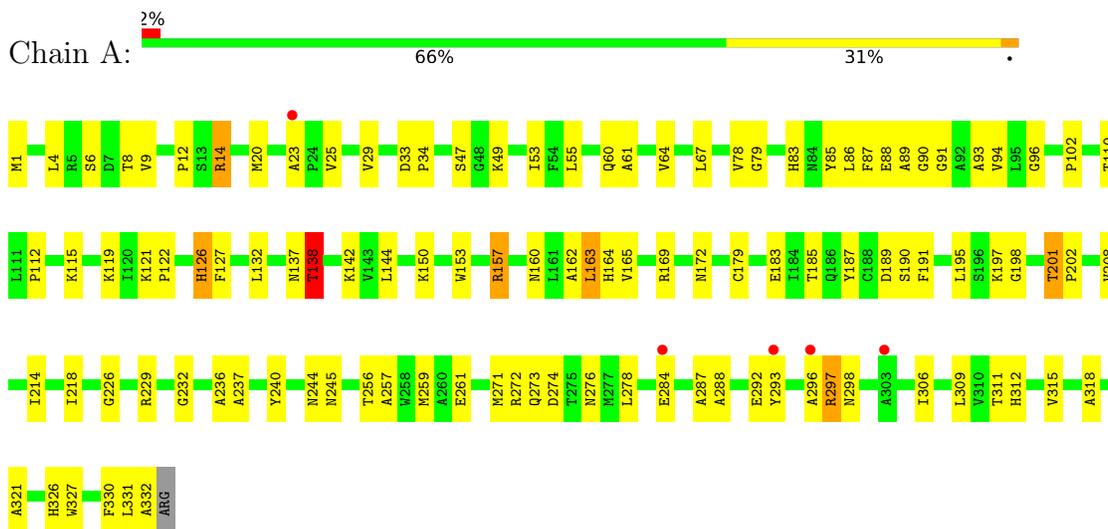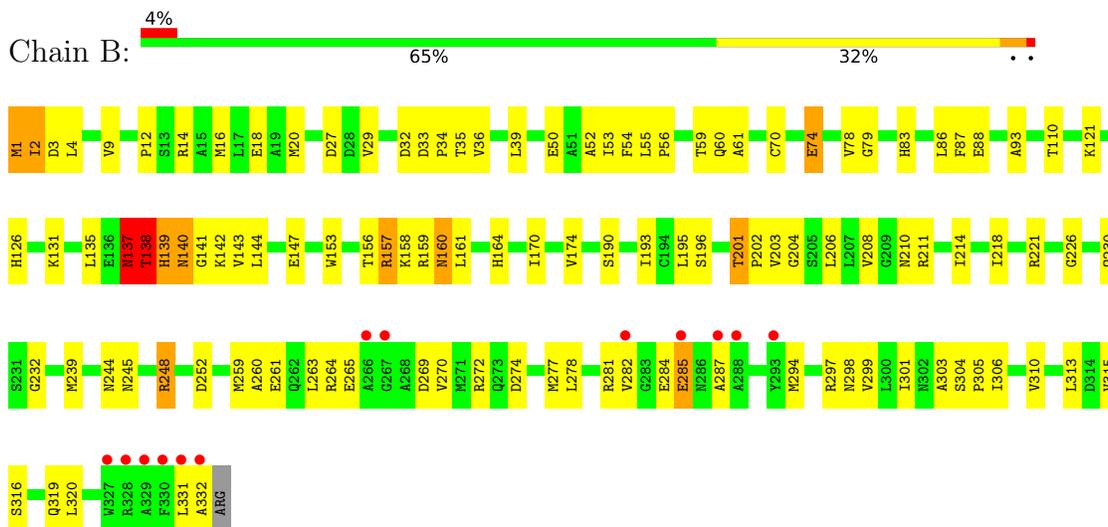
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	76.62Å 100.79Å 176.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.35 – 2.10 38.23 – 2.10	Depositor EDS
% Data completeness (in resolution range)	86.5 (29.35-2.10) 86.9 (38.23-2.10)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.67 (at 2.10Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.209 , 0.279 0.217 , 0.282	Depositor DCC
R_{free} test set	1776 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	20.6	Xtriage
Anisotropy	0.158	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 61.7	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5449	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.70% 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, 0.333 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: EPE, MG, PLR

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	0.49	2/2603 (0.1%)	0.68	3/3525 (0.1%)
1	B	0.70	16/2603 (0.6%)	0.65	0/3525
All	All	0.60	18/5206 (0.3%)	0.67	3/7050 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	126[A]	HIS	CB-CG	-14.19	1.24	1.50
1	A	126[B]	HIS	CB-CG	-14.19	1.24	1.50
1	B	137	ASN	CG-ND2	-8.77	1.10	1.32
1	B	141	GLY	C-O	-8.30	1.10	1.23
1	B	142	LYS	C-O	-8.00	1.08	1.23

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	126[A]	HIS	CA-CB-CG	13.35	136.29	113.60
1	A	126[B]	HIS	CA-CB-CG	13.35	136.29	113.60
1	A	138	THR	N-CA-C	-5.21	96.93	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	137	ASN	Peptide

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	2553	0	2548	94	0
1	B	2554	0	2551	104	0
2	A	2	0	0	0	0
2	B	1	0	0	0	0
3	A	15	0	8	0	0
3	B	15	0	8	0	0
4	A	15	0	18	1	0
4	B	15	0	18	2	0
5	A	144	0	0	7	0
5	B	135	0	0	6	0
All	All	5449	0	5151	197	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 19.

The worst 5 of 197 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:B:138:THR:HG21	5:B:635:HOH:O	1.53	1.06
1:B:60:GLN:HE22	1:B:226:GLY:HA3	1.25	0.98
1:A:60:GLN:HE22	1:A:226:GLY:HA3	1.26	0.97
1:A:25:VAL:HG12	1:A:232:GLY:HA3	1.56	0.87
1:B:138:THR:CG2	5:B:635:HOH:O	2.12	0.87

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	331/333 (99%)	305 (92%)	22 (7%)	4 (1%)	13	8
1	B	331/333 (99%)	301 (91%)	26 (8%)	4 (1%)	13	8
All	All	662/666 (99%)	606 (92%)	48 (7%)	8 (1%)	13	8

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	287	ALA
1	A	89	ALA
1	A	138	THR
1	B	287	ALA
1	B	138	THR

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	261/261 (100%)	252 (97%)	9 (3%)	37	39
1	B	261/261 (100%)	251 (96%)	10 (4%)	33	34
All	All	522/522 (100%)	503 (96%)	19 (4%)	36	36

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	157	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	248	ARG
1	B	285	GLU
1	B	201	THR
1	A	284	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	137	ASN
1	B	172	ASN
1	B	164	HIS
1	B	230	GLN
1	A	230	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 7 ligands modelled in this entry, 3 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
3	PLR	A	403	1	15,15,15	5.52	7 (46%)	20,22,22	2.84	9 (45%)
3	PLR	B	402	1	15,15,15	5.72	6 (40%)	20,22,22	2.92	8 (40%)
4	EPE	B	403	-	15,15,15	1.75	5 (33%)	18,20,20	1.68	4 (22%)
4	EPE	A	404	-	15,15,15	1.37	1 (6%)	18,20,20	1.44	2 (11%)

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	PLR	A	403	1	-	2/6/6/6	0/1/1/1
3	PLR	B	402	1	-	3/6/6/6	0/1/1/1
4	EPE	B	403	-	-	3/9/19/19	0/1/1/1
4	EPE	A	404	-	-	2/9/19/19	0/1/1/1

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	402	PLR	C5-C4	17.98	1.60	1.40
3	A	403	PLR	C5-C4	17.22	1.59	1.40
3	B	402	PLR	C3-C2	9.65	1.50	1.40
3	A	403	PLR	C3-C2	9.29	1.50	1.40
3	B	402	PLR	C2-N1	4.78	1.42	1.33

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	402	PLR	C2A-C2-C3	6.63	129.07	120.89
3	A	403	PLR	C2A-C2-C3	6.44	128.85	120.89
3	B	402	PLR	C4A-C4-C5	5.89	127.00	120.94
3	A	403	PLR	C4A-C4-C5	5.42	126.52	120.94
3	A	403	PLR	C6-N1-C2	4.55	127.60	119.17

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	403	PLR	C4-C5-C5A-O4P
3	A	403	PLR	C6-C5-C5A-O4P

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	B	402	PLR	C4-C5-C5A-O4P
3	B	402	PLR	C6-C5-C5A-O4P
4	A	404	EPE	S-C10-C9-N1

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	403	EPE	2	0
4	A	404	EPE	1	0

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	332/333 (99%)	-0.24	5 (1%)	73	77	10, 25, 46, 66	0
1	B	332/333 (99%)	-0.15	13 (3%)	39	45	9, 25, 61, 84	0
All	All	664/666 (99%)	-0.19	18 (2%)	54	60	9, 25, 51, 84	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	328	ARG	3.6
1	B	330	PHE	3.2
1	B	266	ALA	3.0
1	A	293	TYR	3.0
1	A	23	ALA	2.9

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	EPE	B	403	15/15	0.82	0.20	69,76,80,80	0
4	EPE	A	404	15/15	0.91	0.16	33,42,52,56	0
3	PLR	B	402	15/15	0.97	0.13	17,25,33,34	0
2	MG	A	401	1/1	0.97	0.11	20,20,20,20	0
3	PLR	A	403	15/15	0.97	0.11	11,24,40,42	0
2	MG	A	402	1/1	0.98	0.16	23,23,23,23	0
2	MG	B	401	1/1	0.98	0.13	21,21,21,21	0

6.5 Other polymers ⓘ

There are no such residues in this entry.