wwPDB X-ray Structure Validation Summary Report (i)

Oct 8, 2023 – 07:57 AM EDT

PDB ID	:	6E4B
Title	:	The crystal structure of a putative alpha-ribazole-5'-P phosphatase from Escherichia coli str. K-12 substr. MG1655
Authors	:	Tan, K.; Maltseva, N.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on	:	2018-07-17
Resolution	:	2.55 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:

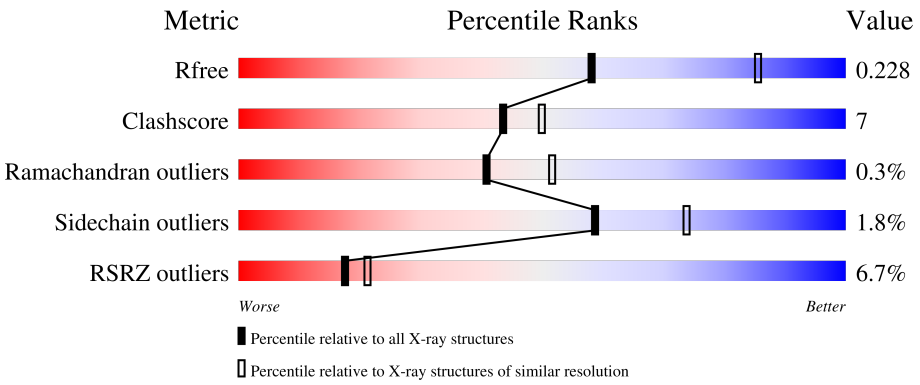
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.55 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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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	227	77% 12% 11%
1	B	227	% 78% 11% 11%
1	C	227	3% 75% 14% 11%
1	D	227	20% 69% 18% 13%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6334 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 Adenosylcobalamin/alpha-ribazole phosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	203	1640	1037	301	292	10	0	1	0
1	B	203	1638	1034	301	294	9	0	0	0
1	C	203	1579	1004	291	276	8	0	0	0
1	D	198	1367	863	254	242	8	0	0	0

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-23	MET	-	initiating methionine	UNP P52086
A	-22	HIS	-	expression tag	UNP P52086
A	-21	HIS	-	expression tag	UNP P52086
A	-20	HIS	-	expression tag	UNP P52086
A	-19	HIS	-	expression tag	UNP P52086
A	-18	HIS	-	expression tag	UNP P52086
A	-17	HIS	-	expression tag	UNP P52086
A	-16	SER	-	expression tag	UNP P52086
A	-15	SER	-	expression tag	UNP P52086
A	-14	GLY	-	expression tag	UNP P52086
A	-13	VAL	-	expression tag	UNP P52086
A	-12	ASP	-	expression tag	UNP P52086
A	-11	LEU	-	expression tag	UNP P52086
A	-10	GLY	-	expression tag	UNP P52086
A	-9	THR	-	expression tag	UNP P52086
A	-8	GLU	-	expression tag	UNP P52086
A	-7	ASN	-	expression tag	UNP P52086
A	-6	LEU	-	expression tag	UNP P52086
A	-5	TYR	-	expression tag	UNP P52086
A	-4	PHE	-	expression tag	UNP P52086
A	-3	GLN	-	expression tag	UNP P52086

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP P52086
A	-1	ASN	-	expression tag	UNP P52086
A	0	ALA	-	expression tag	UNP P52086
B	-23	MET	-	initiating methionine	UNP P52086
B	-22	HIS	-	expression tag	UNP P52086
B	-21	HIS	-	expression tag	UNP P52086
B	-20	HIS	-	expression tag	UNP P52086
B	-19	HIS	-	expression tag	UNP P52086
B	-18	HIS	-	expression tag	UNP P52086
B	-17	HIS	-	expression tag	UNP P52086
B	-16	SER	-	expression tag	UNP P52086
B	-15	SER	-	expression tag	UNP P52086
B	-14	GLY	-	expression tag	UNP P52086
B	-13	VAL	-	expression tag	UNP P52086
B	-12	ASP	-	expression tag	UNP P52086
B	-11	LEU	-	expression tag	UNP P52086
B	-10	GLY	-	expression tag	UNP P52086
B	-9	THR	-	expression tag	UNP P52086
B	-8	GLU	-	expression tag	UNP P52086
B	-7	ASN	-	expression tag	UNP P52086
B	-6	LEU	-	expression tag	UNP P52086
B	-5	TYR	-	expression tag	UNP P52086
B	-4	PHE	-	expression tag	UNP P52086
B	-3	GLN	-	expression tag	UNP P52086
B	-2	SER	-	expression tag	UNP P52086
B	-1	ASN	-	expression tag	UNP P52086
B	0	ALA	-	expression tag	UNP P52086
C	-23	MET	-	initiating methionine	UNP P52086
C	-22	HIS	-	expression tag	UNP P52086
C	-21	HIS	-	expression tag	UNP P52086
C	-20	HIS	-	expression tag	UNP P52086
C	-19	HIS	-	expression tag	UNP P52086
C	-18	HIS	-	expression tag	UNP P52086
C	-17	HIS	-	expression tag	UNP P52086
C	-16	SER	-	expression tag	UNP P52086
C	-15	SER	-	expression tag	UNP P52086
C	-14	GLY	-	expression tag	UNP P52086
C	-13	VAL	-	expression tag	UNP P52086
C	-12	ASP	-	expression tag	UNP P52086
C	-11	LEU	-	expression tag	UNP P52086
C	-10	GLY	-	expression tag	UNP P52086
C	-9	THR	-	expression tag	UNP P52086

Continued on next page...

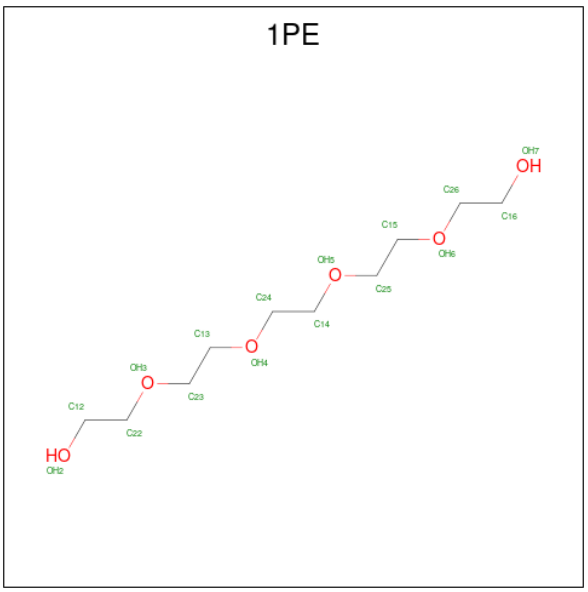
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-8	GLU	-	expression tag	UNP P52086
C	-7	ASN	-	expression tag	UNP P52086
C	-6	LEU	-	expression tag	UNP P52086
C	-5	TYR	-	expression tag	UNP P52086
C	-4	PHE	-	expression tag	UNP P52086
C	-3	GLN	-	expression tag	UNP P52086
C	-2	SER	-	expression tag	UNP P52086
C	-1	ASN	-	expression tag	UNP P52086
C	0	ALA	-	expression tag	UNP P52086
D	-23	MET	-	initiating methionine	UNP P52086
D	-22	HIS	-	expression tag	UNP P52086
D	-21	HIS	-	expression tag	UNP P52086
D	-20	HIS	-	expression tag	UNP P52086
D	-19	HIS	-	expression tag	UNP P52086
D	-18	HIS	-	expression tag	UNP P52086
D	-17	HIS	-	expression tag	UNP P52086
D	-16	SER	-	expression tag	UNP P52086
D	-15	SER	-	expression tag	UNP P52086
D	-14	GLY	-	expression tag	UNP P52086
D	-13	VAL	-	expression tag	UNP P52086
D	-12	ASP	-	expression tag	UNP P52086
D	-11	LEU	-	expression tag	UNP P52086
D	-10	GLY	-	expression tag	UNP P52086
D	-9	THR	-	expression tag	UNP P52086
D	-8	GLU	-	expression tag	UNP P52086
D	-7	ASN	-	expression tag	UNP P52086
D	-6	LEU	-	expression tag	UNP P52086
D	-5	TYR	-	expression tag	UNP P52086
D	-4	PHE	-	expression tag	UNP P52086
D	-3	GLN	-	expression tag	UNP P52086
D	-2	SER	-	expression tag	UNP P52086
D	-1	ASN	-	expression tag	UNP P52086
D	0	ALA	-	expression tag	UNP P52086

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

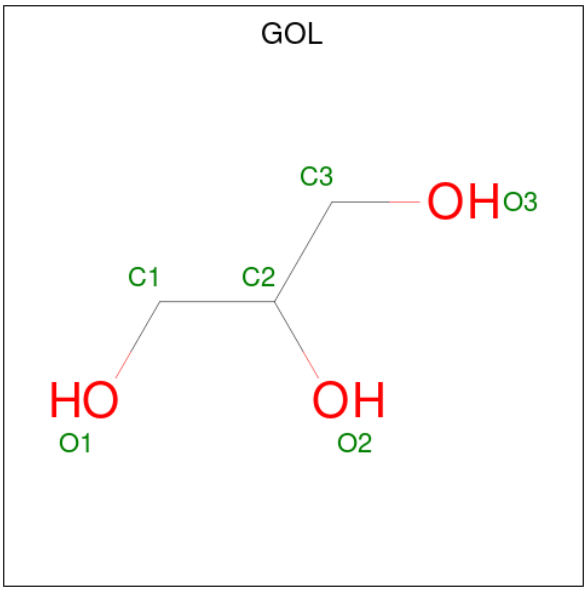
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
2	A	2	2	2	0	0
			Total	Cl		
2	B	1	1	1	0	0

- Molecule 3 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 16 10 6	0	0
3	A	1	Total C O 16 10 6	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 6 3 3	0	0

- Molecule 5 is water.

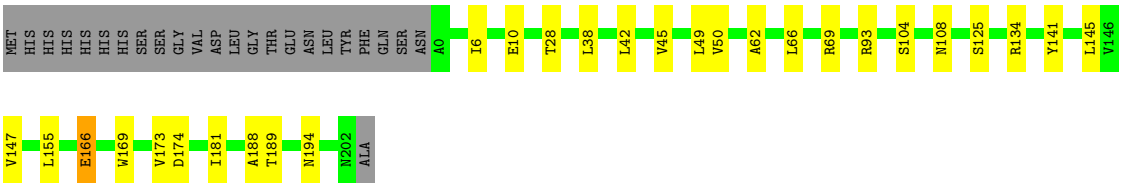
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	48	48	48	0	0
			Total	O		
5	B	16	16	16	0	0
			Total	O		
5	C	3	3	3	0	0
			Total	O		
5	D	2	2	2	0	0

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.

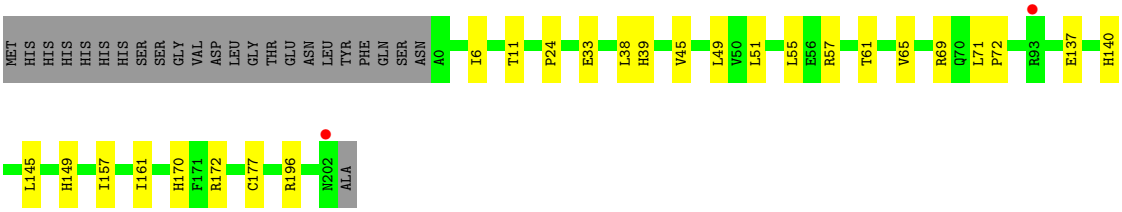
• Molecule 1: Adenosylcobalamin/alpha-ribazole phosphatase

Chain A:



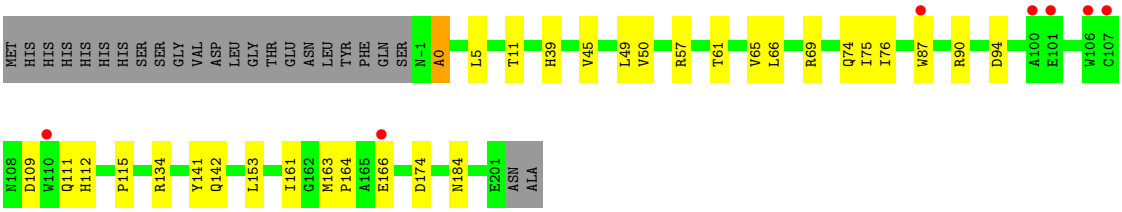
• Molecule 1: Adenosylcobalamin/alpha-ribazole phosphatase

Chain B:



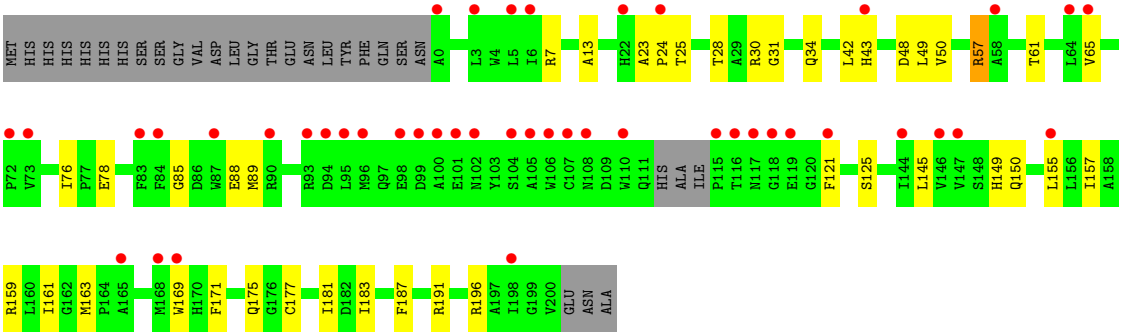
• Molecule 1: Adenosylcobalamin/alpha-ribazole phosphatase

Chain C:



• Molecule 1: Adenosylcobalamin/alpha-ribazole phosphatase

Chain D:


MET HIS HIS HIS HIS HIS HIS SER SER GLY VAL ASP LEU GLY THR GLU ASN LEU TYR PHE GLN SER ASN A0 L3 W4 L5 I6 R7 A13 H22 A23 P24 T25 T28 A29 R30 G31 Q34 L42 H43 D48 L49 V50 R57 A58 T61 L64 V65
P72 V73 I76 P77 E78 F83 F84 G85 D86 W87 E88 M89 R90 R93 D94 L95 M96 Q97 E98 D99 A100 E101 N102 Y103 S104 A105 W106 C107 N108 D109 W110 Q111 HIS ALA ILE P115 T116 N117 G118 E119 G120 F121 S125 I144 L145 V146 V147 S148 H149 Q150 L155 L156 I157 A158
R159 L160 I161 G162 M163 P164 A165 M168 W169 H170 F171 Q175 G176 C177 I181 D182 I183 F187 R191 R196 A197 I198 G199 V200 GLU ASN ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	114.11Å 53.77Å 199.13Å 90.00° 98.09° 90.00°	Depositor
Resolution (Å)	46.17 – 2.55 48.55 – 2.55	Depositor EDS
% Data completeness (in resolution range)	84.2 (46.17-2.55) 84.2 (48.55-2.55)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.31 (at 2.54Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.193 , 0.228 0.193 , 0.228	Depositor DCC
R_{free} test set	1663 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	36.9	Xtriage
Anisotropy	0.223	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 65.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6334	wwPDB-VP
Average B, all atoms (Å^2)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.78% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: 1PE, GOL, CL

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	0.24	0/1683	0.41	0/2283
1	B	0.24	0/1678	0.40	0/2277
1	C	0.24	0/1618	0.40	0/2201
1	D	0.24	0/1394	0.41	0/1909
All	All	0.24	0/6373	0.41	0/8670

There are no bond length outliers.

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	1640	0	1592	21	0
1	B	1638	0	1585	18	0
1	C	1579	0	1499	19	0
1	D	1367	0	1181	26	0
2	A	2	0	0	0	0
2	B	1	0	0	1	0
3	A	32	0	44	3	0
4	B	6	0	8	0	0
5	A	48	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	16	0	0	2	0
5	C	3	0	0	0	0
5	D	2	0	0	1	0
All	All	6334	0	5909	81	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 7.

The worst 5 of 81 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:C:111:GLN:NE2	1:C:166:GLU:HB3	1.90	0.86
1:D:159:ARG:NH1	1:D:163:MET:O	2.13	0.81
1:A:66:LEU:HD23	1:A:69:ARG:NE	1.96	0.80
1:A:66:LEU:HD23	1:A:69:ARG:HE	1.52	0.73
1:D:13:ALA:H	1:D:25:THR:HG21	1.58	0.68

There are no symmetry-related clashes.

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	202/227 (89%)	193 (96%)	9 (4%)	0	100	100
1	B	201/227 (88%)	194 (96%)	7 (4%)	0	100	100
1	C	201/227 (88%)	186 (92%)	14 (7%)	1 (0%)	29	40
1	D	194/227 (86%)	184 (95%)	9 (5%)	1 (0%)	29	40
All	All	798/908 (88%)	757 (95%)	39 (5%)	2 (0%)	41	51

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	49	LEU
1	C	0	ALA

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	172/194 (89%)	169 (98%)	3 (2%)	60	75
1	B	172/194 (89%)	171 (99%)	1 (1%)	86	92
1	C	155/194 (80%)	153 (99%)	2 (1%)	69	80
1	D	109/194 (56%)	104 (95%)	5 (5%)	27	36
All	All	608/776 (78%)	597 (98%)	11 (2%)	59	74

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	43	HIS
1	D	57	ARG
1	D	171	PHE
1	D	89	MET
1	C	94	ASP

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

Mol	Chain	Res	Type
1	A	194	ASN
1	C	111	GLN
1	D	175	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)

Of 6 ligands modelled in this entry, 3 are monoatomic - leaving 3 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	1PE	A	304	-	15,15,15	0.70	0	14,14,14	0.22	0
3	1PE	A	303	-	15,15,15	0.69	0	14,14,14	0.22	0
4	GOL	B	302	-	5,5,5	0.98	0	5,5,5	0.89	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	1PE	A	304	-	-	8/13/13/13	-
3	1PE	A	303	-	-	6/13/13/13	-
4	GOL	B	302	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	302	GOL	O1-C1-C2-O2
4	B	302	GOL	O1-C1-C2-C3
3	A	303	1PE	OH4-C13-C23-OH3
3	A	303	1PE	OH5-C14-C24-OH4
3	A	304	1PE	OH6-C15-C25-OH5

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	304	1PE	2	0
3	A	303	1PE	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	203/227 (89%)	-0.08	0 100 100	13, 25, 52, 86	0
1	B	203/227 (89%)	0.14	2 (0%) 82 86	20, 44, 69, 107	0
1	C	203/227 (89%)	0.34	7 (3%) 45 52	37, 63, 112, 133	0
1	D	198/227 (87%)	1.04	45 (22%) 0 0	52, 89, 133, 157	0
All	All	807/908 (88%)	0.35	54 (6%) 17 21	13, 52, 115, 157	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	100	ALA	8.7
1	D	106	TRP	7.1
1	D	84	PHE	6.3
1	D	87	TRP	5.9
1	D	118	GLY	5.5

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
4	GOL	B	302	6/6	0.90	0.22	44,54,63,63	0
3	1PE	A	304	16/16	0.93	0.17	39,58,74,75	0
3	1PE	A	303	16/16	0.95	0.19	40,49,62,72	0
2	CL	B	301	1/1	0.97	0.08	46,46,46,46	0
2	CL	A	301	1/1	1.00	0.13	26,26,26,26	0
2	CL	A	302	1/1	1.00	0.09	37,37,37,37	0

6.5 Other polymers (i)

There are no such residues in this entry.