W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Nov 7, 2023 – 08:01 AM EST

PDB ID	:	6CJB
Title	:	Crystal structure of Cystathionine beta-lyase from Legionella pneumophila Philadelphia 1 covalently bound to Pyridoxal phosphate
Authors	:	Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on	:	2018-02-26
Resolution	:	1.75 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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance ⓘ

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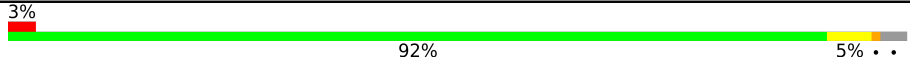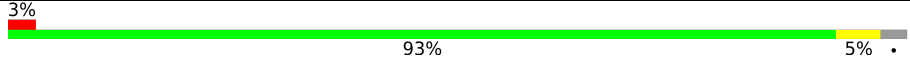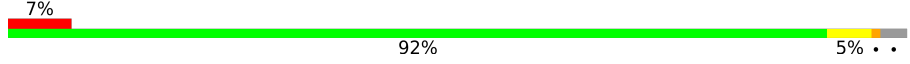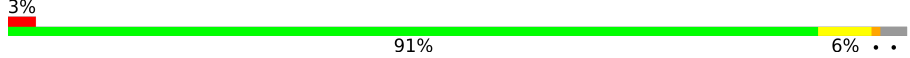*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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.
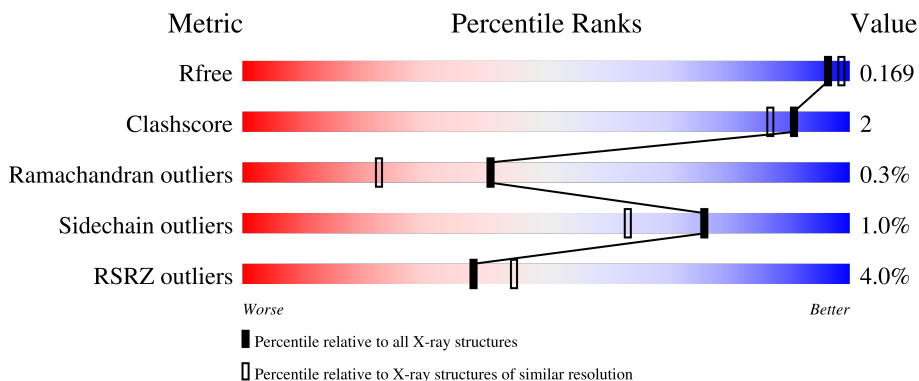
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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	391	3% / 92% / 5% · ·
1	B	391	3% / 93% / 5% ·
1	C	391	7% / 92% / 5% · ·
1	D	391	3% / 91% / 6% · ·

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 13294 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 Cystathionine beta-lyase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	381	Total	C	N	O	P	S	0	20	0
			2997	1915	506	561	1	14			
1	B	381	Total	C	N	O	P	S	0	22	0
			3020	1932	507	566	1	14			
1	C	381	Total	C	N	O	P	S	0	13	0
			2956	1882	504	555	1	14			
1	D	381	Total	C	N	O	P	S	0	18	0
			2975	1901	506	553	1	14			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	initiating methionine	UNP Q5ZX43
A	-6	ALA	-	expression tag	UNP Q5ZX43
A	-5	HIS	-	expression tag	UNP Q5ZX43
A	-4	HIS	-	expression tag	UNP Q5ZX43
A	-3	HIS	-	expression tag	UNP Q5ZX43
A	-2	HIS	-	expression tag	UNP Q5ZX43
A	-1	HIS	-	expression tag	UNP Q5ZX43
A	0	HIS	-	expression tag	UNP Q5ZX43
B	-7	MET	-	initiating methionine	UNP Q5ZX43
B	-6	ALA	-	expression tag	UNP Q5ZX43
B	-5	HIS	-	expression tag	UNP Q5ZX43
B	-4	HIS	-	expression tag	UNP Q5ZX43
B	-3	HIS	-	expression tag	UNP Q5ZX43
B	-2	HIS	-	expression tag	UNP Q5ZX43
B	-1	HIS	-	expression tag	UNP Q5ZX43
B	0	HIS	-	expression tag	UNP Q5ZX43
C	-7	MET	-	initiating methionine	UNP Q5ZX43
C	-6	ALA	-	expression tag	UNP Q5ZX43
C	-5	HIS	-	expression tag	UNP Q5ZX43
C	-4	HIS	-	expression tag	UNP Q5ZX43
C	-3	HIS	-	expression tag	UNP Q5ZX43

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	HIS	-	expression tag	UNP Q5ZX43
C	-1	HIS	-	expression tag	UNP Q5ZX43
C	0	HIS	-	expression tag	UNP Q5ZX43
D	-7	MET	-	initiating methionine	UNP Q5ZX43
D	-6	ALA	-	expression tag	UNP Q5ZX43
D	-5	HIS	-	expression tag	UNP Q5ZX43
D	-4	HIS	-	expression tag	UNP Q5ZX43
D	-3	HIS	-	expression tag	UNP Q5ZX43
D	-2	HIS	-	expression tag	UNP Q5ZX43
D	-1	HIS	-	expression tag	UNP Q5ZX43
D	0	HIS	-	expression tag	UNP Q5ZX43

- Molecule 2 is FORMIC ACID (three-letter code: FMT) (formula: CH_2O_2).
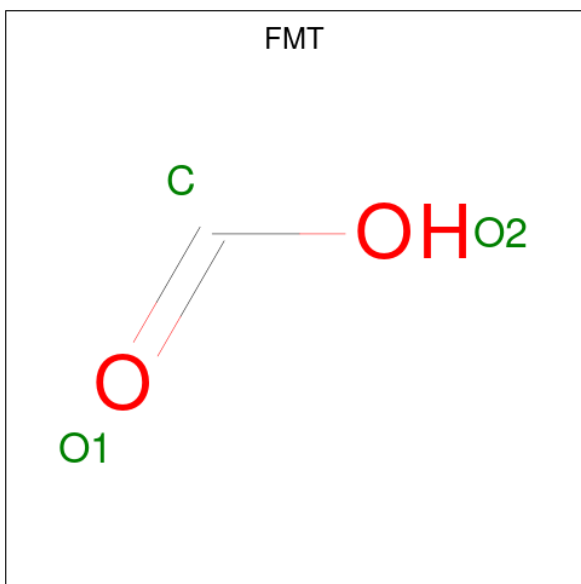
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 3	C 1	O 2	0	0
2	B	1	Total 3	C 1	O 2	0	0
2	C	1	Total 3	C 1	O 2	0	0
2	D	1	Total 3	C 1	O 2	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	322	Total 326	O 326	0	4
4	B	357	Total 364	O 364	0	9
4	C	309	Total 310	O 310	0	2
4	D	325	Total 328	O 328	0	5

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: Cystathionine beta-lyase
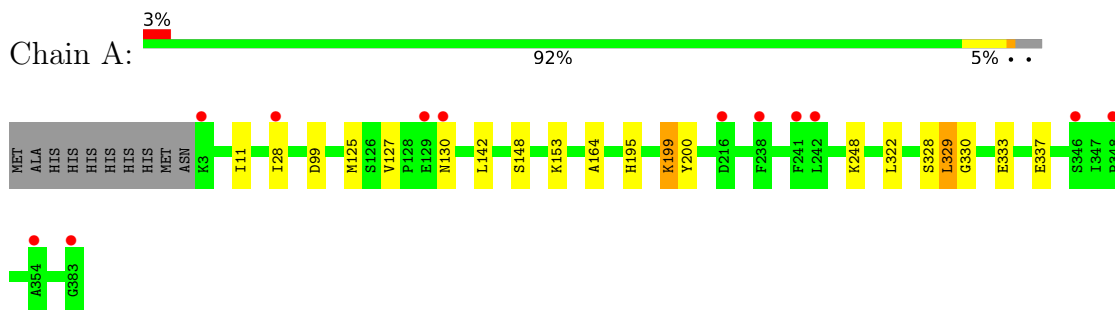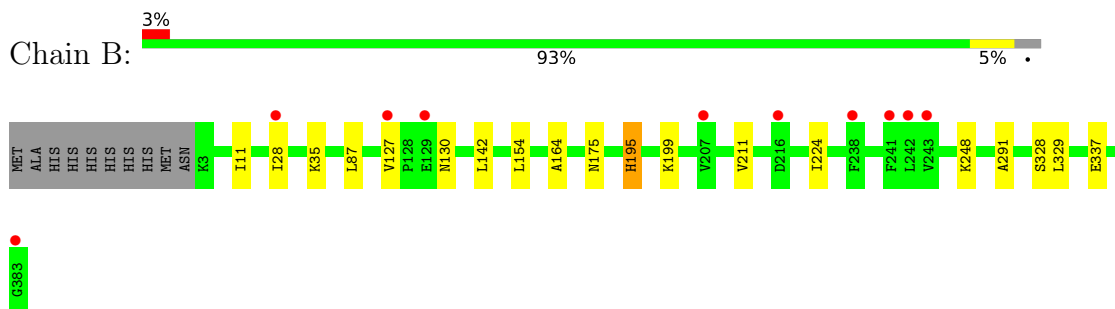
Chain A:

• Molecule 1: Cystathionine beta-lyase

Chain B:

• Molecule 1: Cystathionine beta-lyase

Chain C:

• Molecule 1: Cystathionine beta-lyase

Chain D:

MET ALA HIS HIS HIS HIS HIS MET ASN K3 I11 I28 I81 M97 M125 S126 V127 P128 I131 L142 W143 A164 N168 T171 N175 L188 H195 K199 Y200 V207 V208 V211 V214 G215 D216 I224 P237 F238 F241

L242 V243 K248 K273 L306 F316 S328 L329 G330 E337 G383
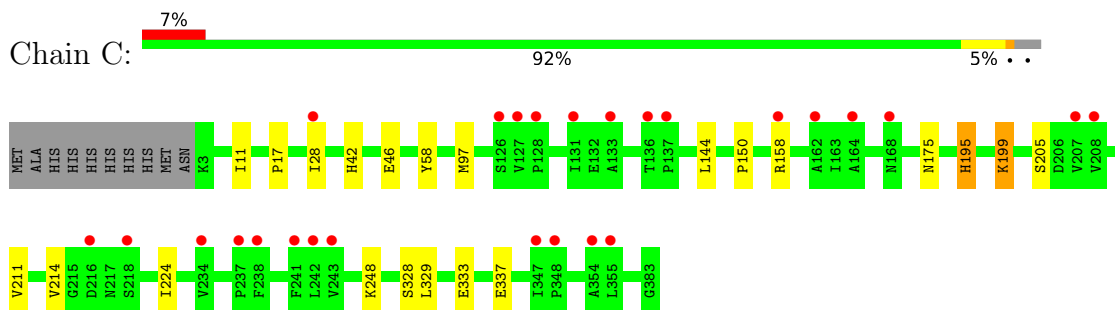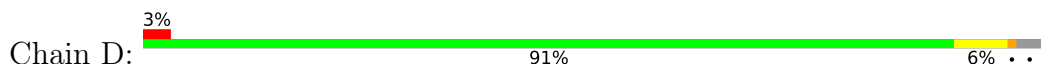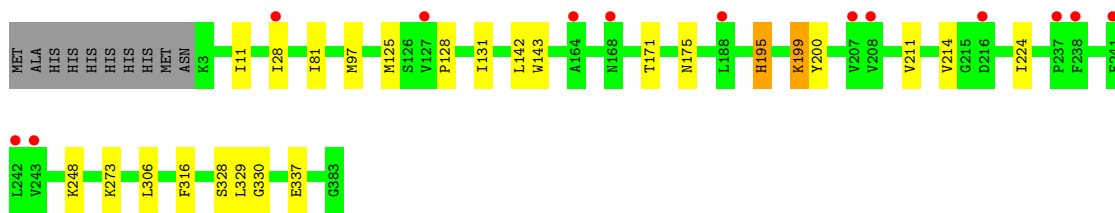
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	96.07Å 97.06Å 176.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.80 – 1.75 46.80 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.80-1.75) 99.7 (46.80-1.75)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.17 (at 1.75Å)	Xtriage
Refinement program	PHENIX (dev_3026)	Depositor
R, R_{free}	0.147 , 0.169 0.147 , 0.169	Depositor DCC
R_{free} test set	1991 reflections (1.20%)	wwPDB-VP
Wilson B-factor (Å^2)	20.5	Xtriage
Anisotropy	0.214	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 50.3	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.009 for k,h,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	13294	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% 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: LLP, FMT, GOL

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	0.46	0/3094	0.61	0/4197
1	B	0.46	0/3123	0.60	0/4235
1	C	0.40	0/3031	0.56	0/4111
1	D	0.38	0/3065	0.58	0/4158
All	All	0.43	0/12313	0.59	0/16701

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	A	0	1
1	B	0	1
1	C	0	1
1	D	0	1
All	All	0	4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	195	HIS	Peptide
1	B	195	HIS	Peptide
1	C	195	HIS	Peptide
1	D	195	HIS	Peptide

5.2   Too-close contacts ⓘ

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	2997	0	3016	10	0
1	B	3020	0	3060	9	0
1	C	2956	0	2949	12	0
1	D	2975	0	2999	13	0
2	A	3	0	1	0	0
2	B	3	0	1	0	0
2	C	3	0	1	0	0
2	D	3	0	1	0	0
3	C	6	0	8	0	0
4	A	326	0	0	0	0
4	B	364	0	0	0	0
4	C	310	0	0	0	0
4	D	328	0	0	0	0
All	All	13294	0	12036	41	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 2.

The worst 5 of 41 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:D:11:ILE:HA	1:D:248[B]:LYS:HG2	1.84	0.59
1:A:99[B]:ASP:OD2	1:A:148:SER:OG	2.23	0.57
1:D:142[B]:LEU:HD22	1:D:171:THR:HG22	1.87	0.56
1:C:42:HIS:NE2	1:C:46[A]:GLU:OE2	2.37	0.54
1:B:142[B]:LEU:HD12	1:B:164:ALA:HB2	1.91	0.53

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	398/391 (102%)	388 (98%)	9 (2%)	1 (0%)	41	22
1	B	400/391 (102%)	390 (98%)	9 (2%)	1 (0%)	41	22
1	C	391/391 (100%)	383 (98%)	7 (2%)	1 (0%)	41	22
1	D	396/391 (101%)	389 (98%)	6 (2%)	1 (0%)	41	22
All	All	1585/1564 (101%)	1550 (98%)	31 (2%)	4 (0%)	41	22

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	328	SER
1	B	328	SER
1	D	328	SER
1	C	328	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	327/327 (100%)	323 (99%)	4 (1%)	71	56
1	B	333/327 (102%)	331 (99%)	2 (1%)	86	79
1	C	318/327 (97%)	314 (99%)	4 (1%)	69	54
1	D	323/327 (99%)	321 (99%)	2 (1%)	86	79
All	All	1301/1308 (100%)	1289 (99%)	12 (1%)	76	67

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	150	PRO
1	C	158	ARG
1	D	337	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	337	GLU
1	A	337	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 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	LLP	B	199	1	23,24,25	1.63	3 (13%)	25,32,34	0.90	1 (4%)
1	LLP	D	199	1	23,24,25	1.55	2 (8%)	25,32,34	0.87	1 (4%)
1	LLP	C	199	1	23,24,25	1.57	3 (13%)	25,32,34	0.85	1 (4%)
1	LLP	A	199	1	23,24,25	1.68	3 (13%)	25,32,34	0.90	1 (4%)

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	LLP	B	199	1	-	4/16/17/19	0/1/1/1
1	LLP	D	199	1	-	3/16/17/19	0/1/1/1
1	LLP	C	199	1	-	4/16/17/19	0/1/1/1
1	LLP	A	199	1	-	4/16/17/19	0/1/1/1

wwPDB X-ray Structure Validation Summary Report

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	199	LLP	C4'-NZ	6.41	1.48	1.27
1	A	199	LLP	C4'-NZ	6.30	1.48	1.27
1	D	199	LLP	C4'-NZ	6.26	1.48	1.27
1	C	199	LLP	C4'-NZ	6.23	1.48	1.27
1	A	199	LLP	CB-CA	3.10	1.57	1.53

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	199	LLP	C4-C4'-NZ	-2.93	110.86	124.31
1	D	199	LLP	C4-C4'-NZ	-2.90	110.97	124.31
1	B	199	LLP	C4-C4'-NZ	-2.87	111.14	124.31
1	A	199	LLP	C4-C4'-NZ	-2.61	112.33	124.31

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	199	LLP	O-C-CA-CB
1	B	199	LLP	O-C-CA-CB
1	C	199	LLP	O-C-CA-CB
1	D	199	LLP	O-C-CA-CB
1	A	199	LLP	C4-C4'-NZ-CE

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	D	199	LLP	1	0
1	C	199	LLP	1	0
1	A	199	LLP	1	0

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

5 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	FMT	B	401	-	2,2,2	0.61	0	1,1,1	0.45	0
2	FMT	A	401	-	2,2,2	0.61	0	1,1,1	0.32	0
3	GOL	C	402	-	5,5,5	0.85	0	5,5,5	1.18	1 (20%)
2	FMT	D	401	-	2,2,2	0.65	0	1,1,1	0.50	0
2	FMT	C	401	-	2,2,2	0.61	0	1,1,1	0.22	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	GOL	C	402	-	-	2/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	402	GOL	C3-C2-C1	-2.05	103.75	111.70

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	402	GOL	C1-C2-C3-O3
3	C	402	GOL	O2-C2-C3-O3

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	380/391 (97%)	-0.04	12 (3%)	47	54	11, 22, 45, 60	0
1	B	380/391 (97%)	-0.10	10 (2%)	56	62	10, 20, 39, 61	0
1	C	380/391 (97%)	0.23	26 (6%)	17	22	14, 28, 55, 75	0
1	D	380/391 (97%)	0.08	13 (3%)	45	51	14, 26, 48, 66	0
All	All	1520/1564 (97%)	0.04	61 (4%)	38	45	10, 24, 47, 75	0

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	242	LEU	4.0
1	C	347	ILE	3.8
1	A	383	GLY	3.7
1	C	207	VAL	3.7
1	C	131	ILE	3.6

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

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
1	LLP	D	199	24/25	0.95	0.12	12,21,29,35	0
1	LLP	C	199	24/25	0.97	0.10	17,23,30,37	0
1	LLP	B	199	24/25	0.97	0.10	11,14,24,34	0
1	LLP	A	199	24/25	0.98	0.10	11,16,27,31	0

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
2	FMT	C	401	3/3	0.86	0.13	32,32,33,39	0
3	GOL	C	402	6/6	0.92	0.20	36,45,51,55	0
2	FMT	D	401	3/3	0.94	0.10	26,26,32,34	0
2	FMT	A	401	3/3	0.94	0.09	25,25,29,31	0
2	FMT	B	401	3/3	0.97	0.10	19,19,27,31	0

6.5   Other polymers ⓘ

There are no such residues in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK