wwPDB X-ray Structure Validation Summary Report (i)

Jun 15, 2020 – 02:29 am BST

PDB ID	:	3LER
Title	:	Crystal Structure of Dihydrodipicolinate Synthase from Campylobacter jejuni subsp. jejuni NCTC 11168
Authors	:	Kim, Y.; Zhou, M.; Hasseman, J.; Anderson, W.F.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on	:	2010-01-15
Resolution	:	1.84 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 following versions of software and data (see references (1)) were used in the production of this report:

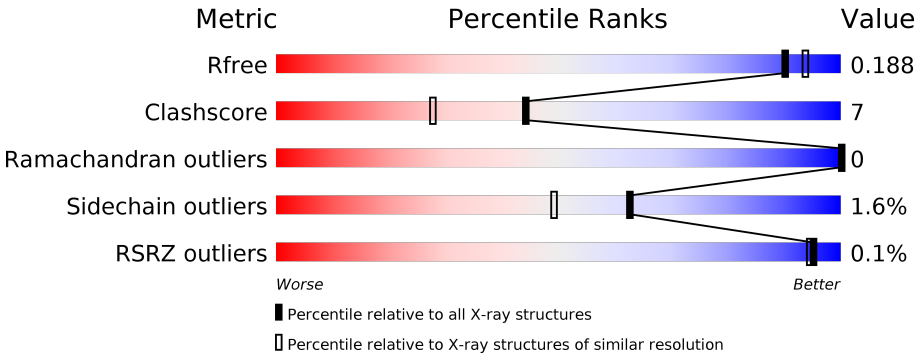
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.84 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	4003 (1.86-1.82)
Clashscore	141614	4233 (1.86-1.82)
Ramachandran outliers	138981	4185 (1.86-1.82)
Sidechain outliers	138945	4186 (1.86-1.82)
RSRZ outliers	127900	3957 (1.86-1.82)

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 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	301	92% 7% •
1	B	301	89% 10% •
1	C	301	83% 14% ••
1	D	301	85% 14% ••

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	PEG	B	303	-	-	X	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 10463 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.

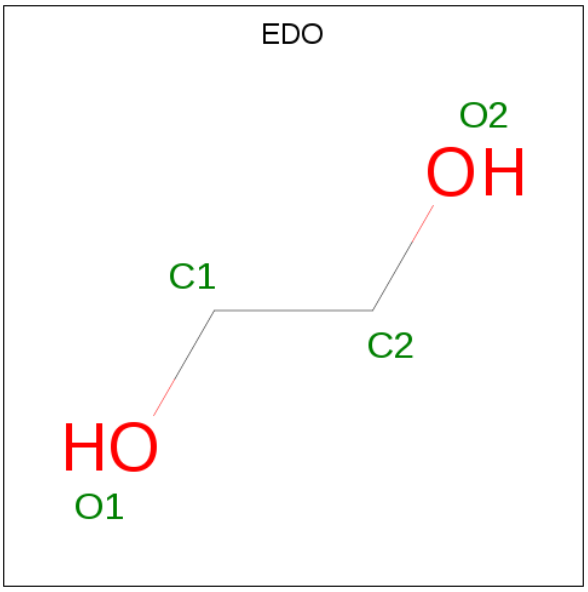
- Molecule 1 is a protein called Dihydrodipicolinate synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	298	Total	C	N	O	S	Se	0	4	0
			2330	1479	387	449	7	8			
1	B	298	Total	C	N	O	S	Se	0	6	0
			2353	1496	394	449	7	7			
1	C	296	Total	C	N	O	S	Se	0	3	0
			2308	1466	387	442	7	6			
1	D	298	Total	C	N	O	S	Se	0	6	0
			2349	1492	389	454	7	7			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP Q9PPB4
A	-1	ASN	-	EXPRESSION TAG	UNP Q9PPB4
A	0	ALA	-	EXPRESSION TAG	UNP Q9PPB4
B	-2	SER	-	EXPRESSION TAG	UNP Q9PPB4
B	-1	ASN	-	EXPRESSION TAG	UNP Q9PPB4
B	0	ALA	-	EXPRESSION TAG	UNP Q9PPB4
C	-2	SER	-	EXPRESSION TAG	UNP Q9PPB4
C	-1	ASN	-	EXPRESSION TAG	UNP Q9PPB4
C	0	ALA	-	EXPRESSION TAG	UNP Q9PPB4
D	-2	SER	-	EXPRESSION TAG	UNP Q9PPB4
D	-1	ASN	-	EXPRESSION TAG	UNP Q9PPB4
D	0	ALA	-	EXPRESSION TAG	UNP Q9PPB4

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

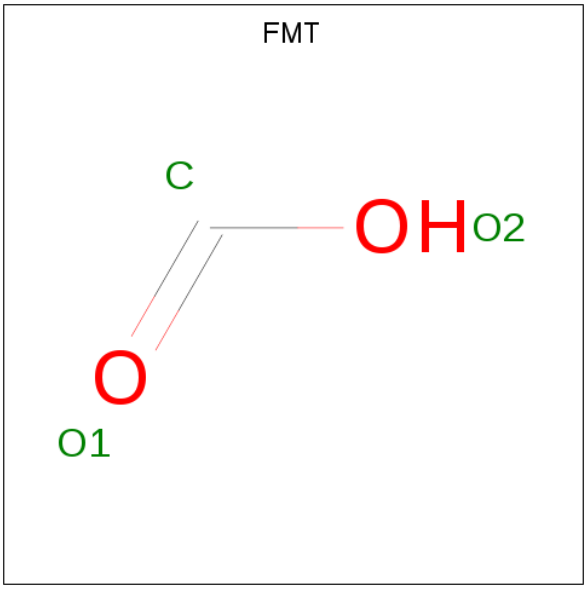


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

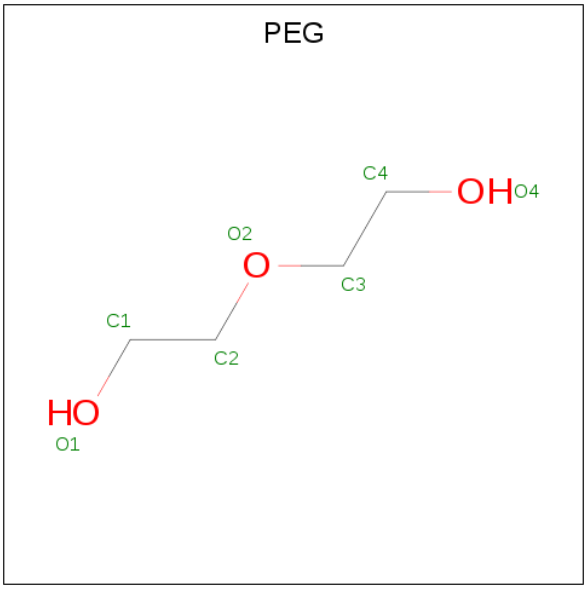
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Mg 1 1	0	0

- Molecule 4 is FORMIC ACID (three-letter code: FMT) (formula: CH_2O_2).



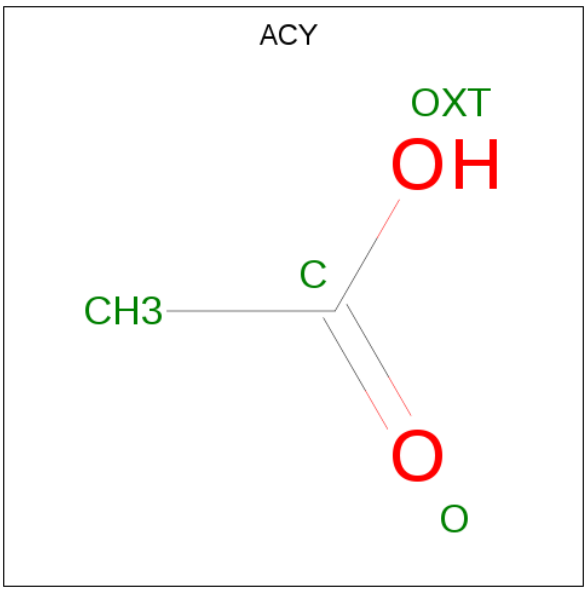
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 3 1 2	0	0

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C O 7 4 3	0	0

- Molecule 6 is ACETIC ACID (three-letter code: ACY) (formula: $C_2H_4O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	D	1	Total C O 4 2 2	0	0

- Molecule 7 is water.

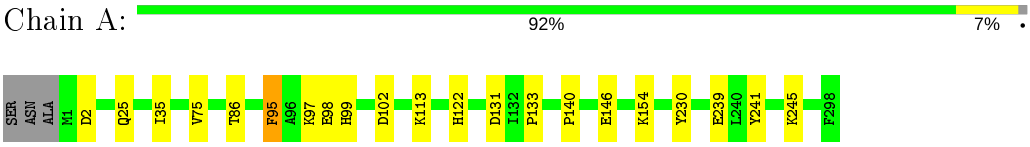
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	279	Total O 279 279	0	0
7	B	294	Total O 294 294	0	0
7	C	226	Total O 226 226	0	0
7	D	301	Total O 301 301	0	0

3 Residue-property plots (i)

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

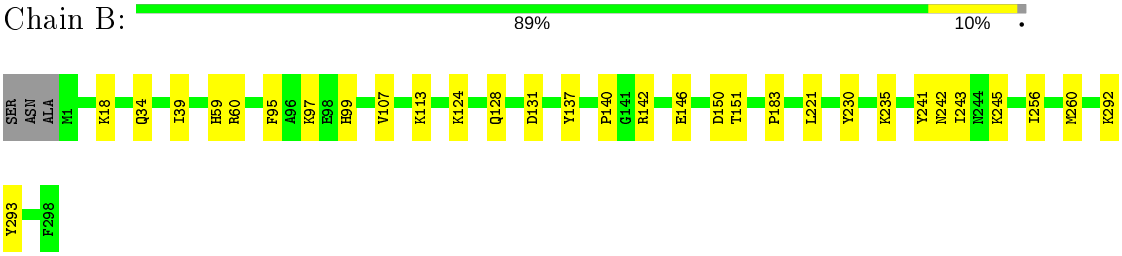
- Molecule 1: Dihydrodipicolinate synthase

Chain A:



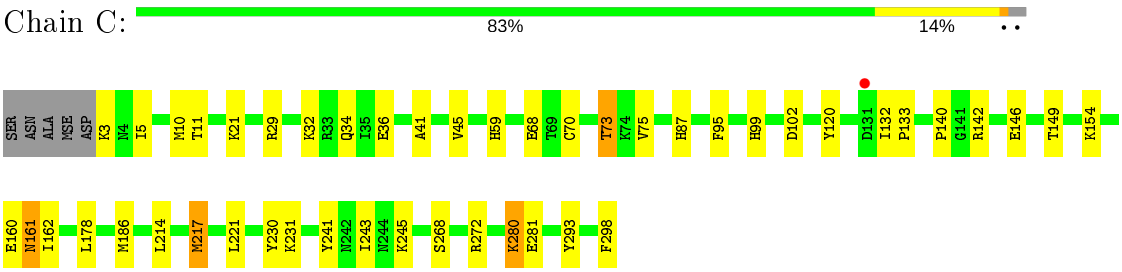
- Molecule 1: Dihydrodipicolinate synthase

Chain B:



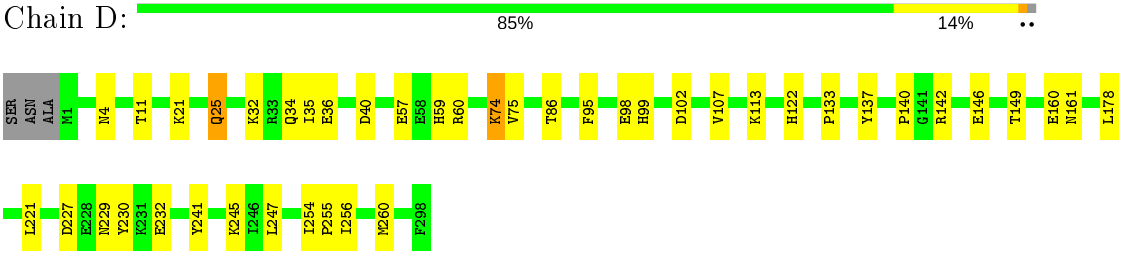
- Molecule 1: Dihydrodipicolinate synthase

Chain C:



- Molecule 1: Dihydrodipicolinate synthase

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	72.19Å 85.30Å 199.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.21 – 1.84 49.79 – 1.84	Depositor EDS
% Data completeness (in resolution range)	99.8 (39.21-1.84) 99.6 (49.79-1.84)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	2.10 (at 1.83Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, R_{free}	0.152 , 0.190 0.150 , 0.188	Depositor DCC
R_{free} test set	5368 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	26.4	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 47.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	10463	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.42% 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: MG, FMT, EDO, KPI, ACY, PEG

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	0.58	0/2349	0.63	0/3164
1	B	0.57	0/2375	0.64	0/3199
1	C	0.52	0/2327	0.64	1/3135 (0.0%)
1	D	0.58	0/2368	0.65	0/3189
All	All	0.56	0/9419	0.64	1/12687 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	217	MSE	CG-SE-CE	-5.79	86.16	98.90

There are no chirality outliers.

There are no planarity outliers.

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	2330	0	2357	22	0
1	B	2353	0	2376	37	0
1	C	2308	0	2340	43	0
1	D	2349	0	2374	36	0
2	A	4	0	6	0	0
2	C	4	0	6	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
4	B	3	0	1	0	0
5	B	7	0	10	4	0
6	D	4	0	3	0	0
7	A	279	0	0	4	1
7	B	294	0	0	5	0
7	C	226	0	0	4	0
7	D	301	0	0	9	1
All	All	10463	0	9473	125	1

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 125 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:113:LYS:HE2	1:C:142[B]:ARG:HH12	1.06	1.13
1:C:73:THR:HG23	1:C:75:VAL:H	1.22	1.01
1:B:113:LYS:HE2	1:D:142:ARG:HD3	1.42	1.00
1:A:230:TYR:CD2	1:D:230:TYR:HD2	1.81	0.98
1:D:95[B]:PHE:HD2	7:D:1073:HOH:O	1.47	0.97

All (1) 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)
7:A:874:HOH:O	7:D:482:HOH:O[3_756]	2.16	0.04

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	299/301 (99%)	293 (98%)	6 (2%)	0	100	100
1	B	301/301 (100%)	294 (98%)	7 (2%)	0	100	100
1	C	296/301 (98%)	289 (98%)	7 (2%)	0	100	100
1	D	301/301 (100%)	294 (98%)	7 (2%)	0	100	100
All	All	1197/1204 (99%)	1170 (98%)	27 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	254/245 (104%)	252 (99%)	2 (1%)	81	75
1	B	256/245 (104%)	255 (100%)	1 (0%)	91	88
1	C	251/245 (102%)	244 (97%)	7 (3%)	43	26
1	D	256/245 (104%)	250 (98%)	6 (2%)	50	34
All	All	1017/980 (104%)	1001 (98%)	16 (2%)	62	49

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

Mol	Chain	Res	Type
1	C	268	SER
1	C	272	ARG
1	D	74	LYS
1	C	161	ASN
1	D	98	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	195	ASN
1	B	242	ASN
1	D	34	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	128	GLN
1	D	59	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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	KPI	C	166	1	10,13,14	1.95	1 (10%)	6,15,17	1.85	3 (50%)
1	KPI	A	166	1	10,13,14	2.19	1 (10%)	6,15,17	1.41	1 (16%)
1	KPI	D	166	1	10,13,14	1.98	1 (10%)	6,15,17	1.93	2 (33%)
1	KPI	B	166	1	10,13,14	2.03	1 (10%)	6,15,17	1.29	1 (16%)

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	KPI	C	166	1	-	1/9/14/16	-
1	KPI	A	166	1	-	1/9/14/16	-
1	KPI	D	166	1	-	1/9/14/16	-
1	KPI	B	166	1	-	1/9/14/16	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	166	KPI	CX1-NZ	6.21	1.47	1.29

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	166	KPI	CX1-NZ	5.98	1.46	1.29
1	D	166	KPI	CX1-NZ	5.81	1.45	1.29
1	C	166	KPI	CX1-NZ	5.74	1.45	1.29

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	166	KPI	CD-CE-NZ	-3.35	104.56	110.66
1	C	166	KPI	CE-NZ-CX1	-2.85	114.11	121.77
1	D	166	KPI	CE-NZ-CX1	-2.45	115.18	121.77
1	A	166	KPI	CE-NZ-CX1	-2.27	115.68	121.77
1	C	166	KPI	CD-CE-NZ	-2.25	106.58	110.66

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	166	KPI	C1-CX1-NZ-CE
1	A	166	KPI	C1-CX1-NZ-CE
1	D	166	KPI	C1-CX1-NZ-CE
1	B	166	KPI	C1-CX1-NZ-CE

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	$\#\|Z\|>2$
4	FMT	B	302	-	0,2,2	0.00	-	0,1,1	0.00	-
2	EDO	A	301	-	3,3,3	0.53	0	2,2,2	0.21	0
2	EDO	C	301	-	3,3,3	0.50	0	2,2,2	0.23	0
6	ACY	D	301	-	1,3,3	1.56	0	0,3,3	0.00	-
5	PEG	B	303	-	6,6,6	0.57	0	5,5,5	0.60	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EDO	A	301	-	-	0/1/1/1	-
2	EDO	C	301	-	-	0/1/1/1	-
5	PEG	B	303	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	303	PEG	O2-C3-C4-O4
5	B	303	PEG	C4-C3-O2-C2

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	301	EDO	1	0
5	B	303	PEG	4	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

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, 95^{th} 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	290/301 (96%)	-0.70	0 100 100	18, 27, 47, 81	0
1	B	290/301 (96%)	-0.61	0 100 100	18, 26, 44, 59	0
1	C	289/301 (96%)	-0.50	1 (0%) 94 93	19, 31, 60, 99	0
1	D	290/301 (96%)	-0.63	0 100 100	18, 26, 44, 87	0
All	All	1159/1204 (96%)	-0.61	1 (0%) 95 94	18, 27, 50, 99	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	131	ASP	2.3

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, 95^{th} 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	KPI	A	166	14/15	0.97	0.08	16,20,26,27	0
1	KPI	C	166	14/15	0.98	0.07	20,23,25,28	0
1	KPI	B	166	14/15	0.98	0.07	16,18,22,22	0
1	KPI	D	166	14/15	0.99	0.07	16,19,23,24	0

6.3 Carbohydrates

There are no carbohydrates 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, 95^{th} 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
5	PEG	B	303	7/7	0.78	0.31	31,46,58,59	0
2	EDO	A	301	4/4	0.81	0.23	58,59,63,64	0
4	FMT	B	302	3/3	0.84	0.18	28,28,48,50	0
2	EDO	C	301	4/4	0.89	0.15	24,40,45,46	0
3	MG	B	301	1/1	0.90	0.09	58,58,58,58	0
6	ACY	D	301	4/4	0.92	0.08	59,60,62,63	0

6.5 Other polymers

There are no such residues in this entry.