wwPDB X-ray Structure Validation Summary Report (i)

Jan 6, 2021 – 06:15 PM GMT

PDB ID	:	6YRC
Title	:	Spectroscopically-validated structure of DtpB from Streptomyces lividans in the ferric state
Authors	:	Lucic, M.; Dworkowski, F.S.N.; Worrall, J.A.R.; Hough, M.A.
Deposited on	:	2020-04-20
Resolution	:	1.99 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 (i)) were used in the production of this report:

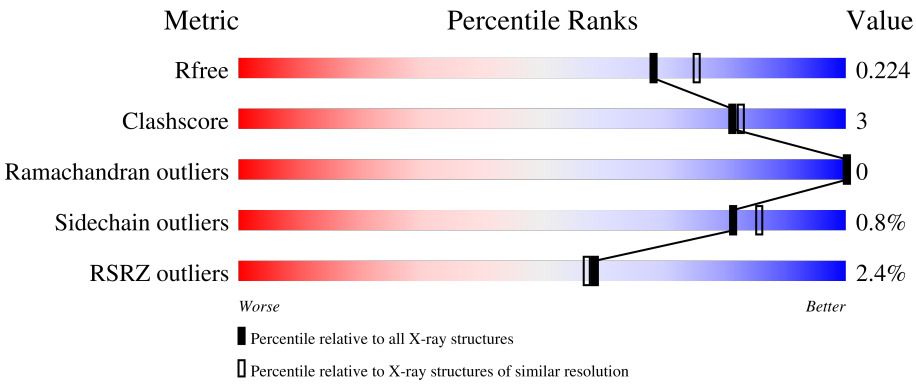
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.16
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.16

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 1.99 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	316	.% 89% 7% •
1	B	316	3% 93% • •
1	C	316	2% 92% 5% •
1	D	316	2% 92% 5% •
1	E	316	3% 95% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	316	3% 90% 8%

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	7PE	A	402	-	-	-	X

2 Entry composition (i)

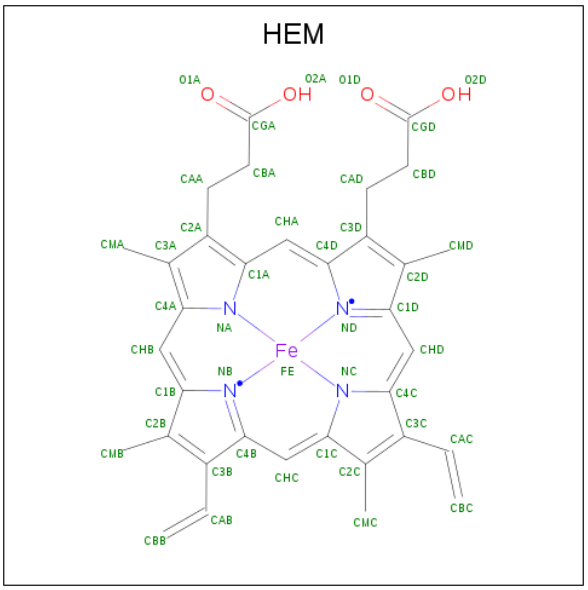
There are 6 unique types of molecules in this entry. The entry contains 15579 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Putative iron-dependent peroxidase.

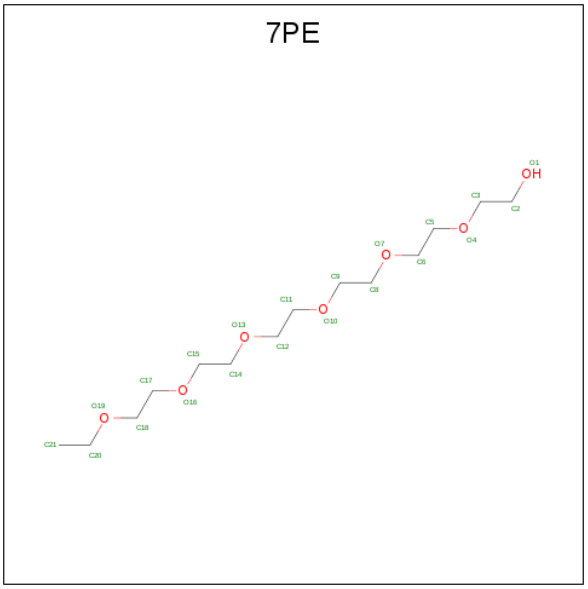
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	306	2372	1488	414	461	9	0	5	0
1	B	306	2328	1466	404	449	9	0	1	0
1	C	306	2330	1468	398	454	10	0	1	0
1	D	306	2348	1474	407	458	9	0	2	0
1	E	306	2335	1471	401	454	9	0	2	0
1	F	308	2349	1476	401	463	9	0	5	0

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Fe	N	O		
2	A	1	43	34	1	4	4	0	0
2	B	1	43	34	1	4	4	0	0
2	C	1	43	34	1	4	4	0	0
2	D	1	43	34	1	4	4	0	0
2	E	1	43	34	1	4	4	0	0
2	F	1	43	34	1	4	4	0	0

- Molecule 3 is 2-(2-(2-(2-(2-(2-ETHOXYETHOXY)ETHOXY)ETHOXY)ETHOXY)ETHOXY)ETHANOL (three-letter code: 7PE) (formula: $C_{14}H_{30}O_7$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	6	4	2	0	0
3	A	1	4	2	2	0	0
3	B	1	5	4	1	0	0
3	B	1	5	3	2	0	0
3	D	1	6	4	2	0	0

Continued on next page...

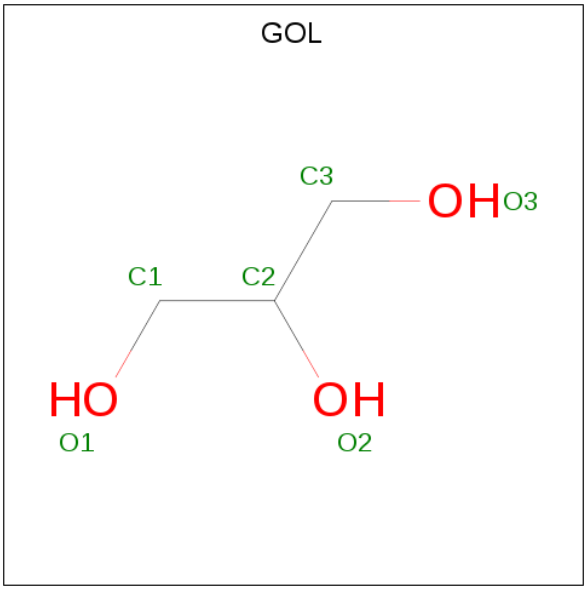
Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	F	1	9	6	3	0	0
			Total	C	O		
3	F	1	6	4	2	0	1
			Total	C	O		
3	F	1	5	3	2	0	0
			Total	C	O		
3	F	1	7	5	2	0	0
			Total	C	O		
3	F	1	5	3	2	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
4	A	3	3	3	0	0
			Total	Mg		
4	C	1	1	1	0	0
			Total	Mg		
4	E	1	1	1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

Continued on next page...

Continued from previous page...

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

- Molecule 6 is water.

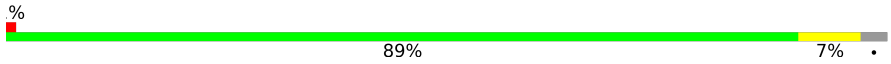
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	216	216	216	0	0
			Total	O		
6	B	147	147	147	0	0
			Total	O		
6	C	205	205	205	0	0
			Total	O		
6	D	198	198	198	0	0
			Total	O		
6	E	189	189	189	0	0
			Total	O		
6	F	187	187	187	0	0

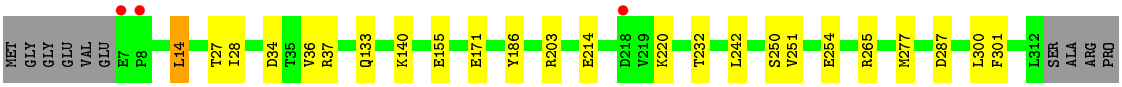
3 Residue-property plots (i)

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: Putative iron-dependent peroxidase

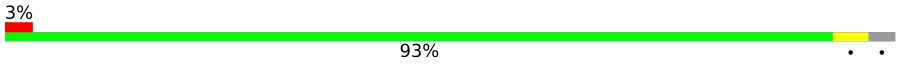
Chain A:

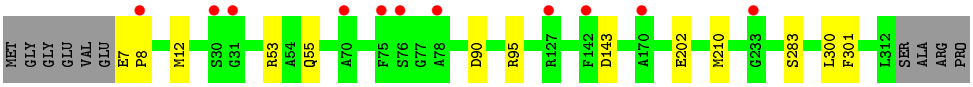




- Molecule 1: Putative iron-dependent peroxidase

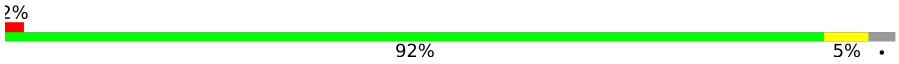
Chain B:





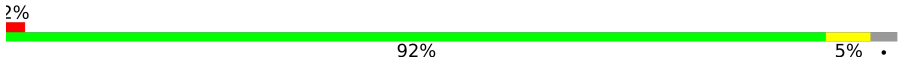
- Molecule 1: Putative iron-dependent peroxidase

Chain C:



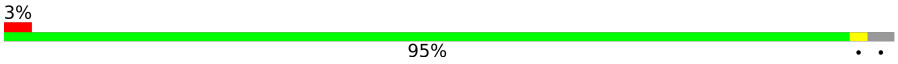
- Molecule 1: Putative iron-dependent peroxidase

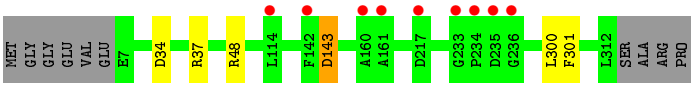
Chain D:



- Molecule 1: Putative iron-dependent peroxidase

Chain E:

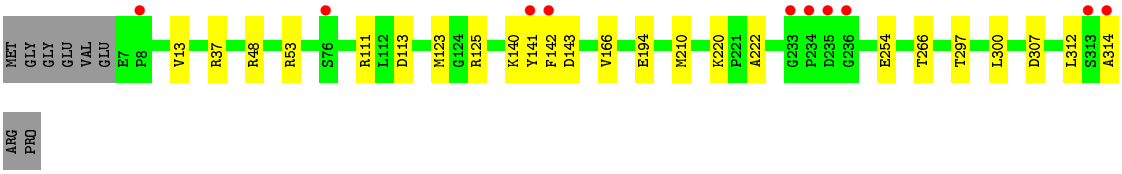




- Molecule 1: Putative iron-dependent peroxidase

Chain F:

3%
90%
8%


MET
GLY
GLY
GLU
VAL
GLU
E7
P8
V13
R37
R48
R53
S76
R111
L112
D113
M123
G124
R125
K140
Y141
F142
D143
V166
E194
M210
K220
P221
A222
G233
P234
D235
G236
E254
T266
T297
L300
D307
L312
S313
A314
ARG
PRO

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.79Å 120.27Å 196.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.05 – 1.99 49.00 – 1.99	Depositor EDS
% Data completeness (in resolution range)	99.0 (49.05-1.99) 99.1 (49.00-1.99)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.06 (at 1.98Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.176 , 0.218 0.185 , 0.224	Depositor DCC
R_{free} test set	6824 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	30.7	Xtriage
Anisotropy	0.029	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 43.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	15579	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.95% 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: GOL, MG, 7PE, HEM

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	0.76	0/2422	0.84	0/3289
1	B	0.71	0/2378	0.81	0/3231
1	C	0.72	0/2380	0.83	0/3233
1	D	0.75	1/2398 (0.0%)	0.85	0/3256
1	E	0.73	0/2385	0.82	0/3241
1	F	0.74	1/2398 (0.0%)	0.81	0/3259
All	All	0.74	2/14361 (0.0%)	0.83	0/19509

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	F	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	199	GLU	CD-OE2	6.57	1.32	1.25
1	F	254	GLU	CD-OE1	-5.49	1.19	1.25

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	F	312	LEU	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	2372	0	2288	16	0
1	B	2328	0	2251	8	0
1	C	2330	0	2251	11	0
1	D	2348	0	2272	13	0
1	E	2335	0	2249	5	0
1	F	2349	0	2233	20	0
2	A	43	0	30	3	0
2	B	43	0	30	2	0
2	C	43	0	30	0	0
2	D	43	0	30	0	0
2	E	43	0	30	0	0
2	F	43	0	30	0	0
3	A	10	0	10	6	0
3	B	10	0	8	1	0
3	D	6	0	6	1	0
3	F	32	0	28	5	0
4	A	3	0	0	0	0
4	C	1	0	0	0	0
4	E	1	0	0	0	0
5	B	12	0	16	0	0
5	C	18	0	24	2	0
5	F	24	0	32	1	0
6	A	216	0	0	4	0
6	B	147	0	0	1	0
6	C	205	0	0	4	0
6	D	198	0	0	5	0
6	E	189	0	0	0	0
6	F	187	0	0	3	0
All	All	15579	0	13848	73	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 3.

The worst 5 of 73 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:214:GLU:OE1	6:C:501:HOH:O	2.05	0.74
1:E:48:ARG:CZ	3:F:410:7PE:H61	2.19	0.73
1:D:214:GLU:OE1	6:D:501:HOH:O	2.13	0.66
1:D:111[B]:ARG:NH1	1:F:142:PHE:HB3	2.10	0.65
1:F:37:ARG:HD2	1:F:314:ALA:HA	1.78	0.65

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	309/316 (98%)	300 (97%)	9 (3%)	0	100	100
1	B	305/316 (96%)	299 (98%)	6 (2%)	0	100	100
1	C	305/316 (96%)	298 (98%)	7 (2%)	0	100	100
1	D	306/316 (97%)	300 (98%)	6 (2%)	0	100	100
1	E	306/316 (97%)	299 (98%)	7 (2%)	0	100	100
1	F	311/316 (98%)	307 (99%)	4 (1%)	0	100	100
All	All	1842/1896 (97%)	1803 (98%)	39 (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	248/252 (98%)	244 (98%)	4 (2%)	62	67
1	B	241/252 (96%)	238 (99%)	3 (1%)	71	76
1	C	243/252 (96%)	242 (100%)	1 (0%)	91	93
1	D	247/252 (98%)	245 (99%)	2 (1%)	81	86
1	E	242/252 (96%)	241 (100%)	1 (0%)	91	93
1	F	241/252 (96%)	238 (99%)	3 (1%)	71	76
All	All	1462/1512 (97%)	1448 (99%)	14 (1%)	81	81

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	283	SER
1	C	127	ARG
1	F	53	ARG
1	B	210	MET
1	E	143	ASP

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

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 30 ligands modelled in this entry, 5 are monoatomic - leaving 25 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

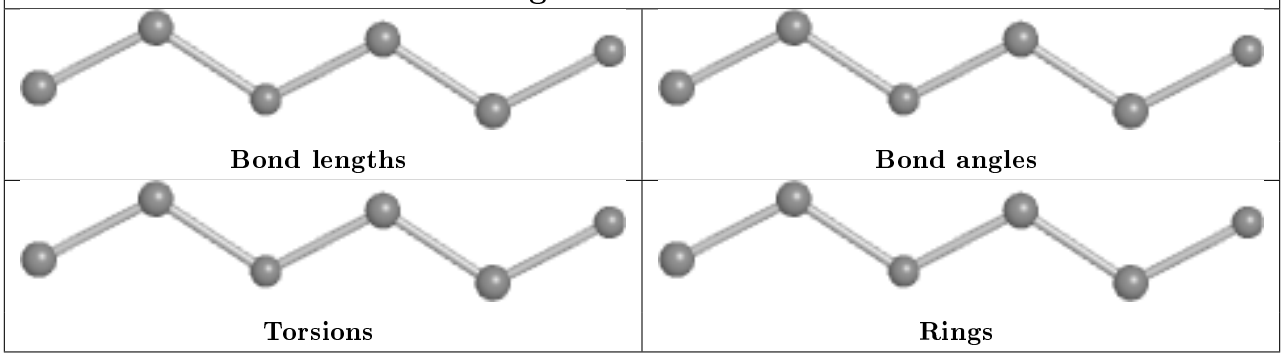
There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand 7PE A 402



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	306/316 (96%)	-0.11	3 (0%)	82	81	22, 28, 45, 69	0
1	B	306/316 (96%)	0.20	11 (3%)	42	42	24, 33, 53, 80	0
1	C	306/316 (96%)	0.00	6 (1%)	65	63	22, 29, 46, 70	0
1	D	306/316 (96%)	0.02	5 (1%)	72	70	23, 29, 48, 67	0
1	E	306/316 (96%)	0.16	9 (2%)	51	50	24, 30, 49, 91	0
1	F	308/316 (97%)	0.02	10 (3%)	47	46	23, 29, 48, 73	0
All	All	1838/1896 (96%)	0.05	44 (2%)	59	57	22, 30, 49, 91	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	234	PRO	6.5
1	C	78	ALA	5.0
1	E	235	ASP	3.8
1	F	142	PHE	3.5
1	F	314	ALA	3.5

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)

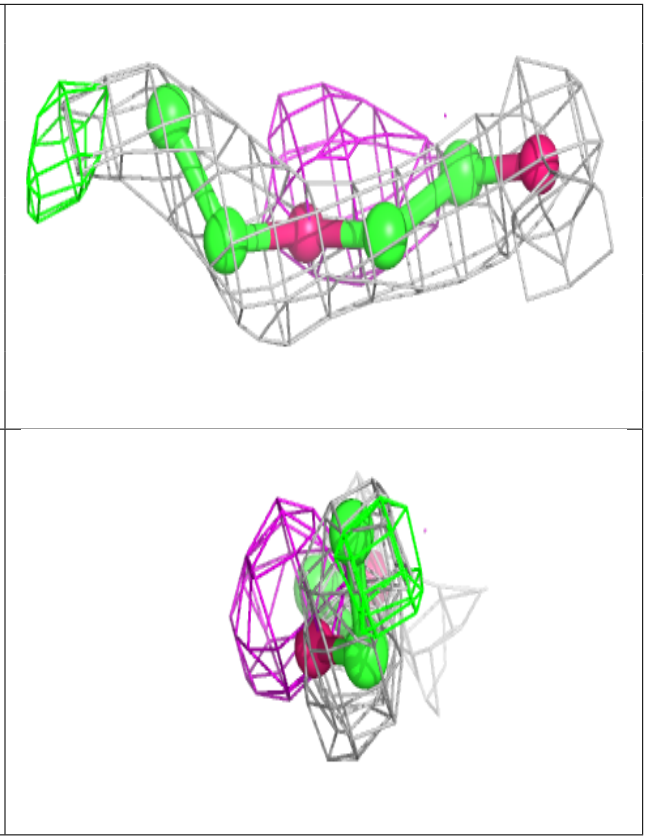
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	7PE	F	409	7/21	0.69	0.22	62,69,74,77	0
4	MG	A	406	1/1	0.72	0.14	56,56,56,56	0
5	GOL	F	407	6/6	0.72	0.21	59,63,67,69	0
3	7PE	A	402	6/21	0.73	0.43	48,52,55,57	0
5	GOL	C	405	6/6	0.74	0.32	48,58,60,62	0
5	GOL	F	403	6/6	0.74	0.17	58,66,68,69	0
5	GOL	B	403	6/6	0.75	0.19	49,59,64,67	0
5	GOL	B	405	6/6	0.79	0.14	54,58,59,60	0
3	7PE	D	402	6/21	0.80	0.33	49,56,56,59	0
5	GOL	F	404	6/6	0.80	0.30	59,62,64,65	0
3	7PE	F	402	9/21	0.80	0.37	39,57,65,72	0
5	GOL	C	403	6/6	0.83	0.18	47,61,62,68	0
5	GOL	C	402	6/6	0.84	0.13	44,52,54,63	0
3	7PE	B	404	5/21	0.84	0.19	49,56,59,60	0
5	GOL	F	406	6/6	0.85	0.20	38,60,65,67	0
3	7PE	F	408	5/21	0.86	0.28	51,54,56,57	0
3	7PE	B	402	5/21	0.88	0.17	50,53,56,58	0
3	7PE	F	410	5/21	0.90	0.31	38,46,49,50	0
3	7PE	A	403	4/21	0.91	0.26	42,42,43,45	0
3	7PE	F	405[A]	6/21	0.92	0.28	14,21,25,25	6
4	MG	E	402	1/1	0.94	0.03	54,54,54,54	0
4	MG	A	405	1/1	0.96	0.04	41,41,41,41	0
2	HEM	C	401	43/43	0.97	0.11	21,25,28,29	0
2	HEM	A	401	43/43	0.97	0.10	21,24,26,32	0
2	HEM	B	401	43/43	0.97	0.10	23,25,28,30	0
4	MG	C	404	1/1	0.97	0.05	30,30,30,30	0
2	HEM	F	401	43/43	0.97	0.10	21,24,26,27	0
2	HEM	E	401	43/43	0.97	0.09	22,25,28,29	0
2	HEM	D	401	43/43	0.98	0.10	21,23,26,27	0
4	MG	A	404	1/1	0.98	0.04	27,27,27,27	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 7PE A 402:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.