Full wwPDB X-ray Structure Validation Report (i)

Sep 6, 2023 – 11:39 PM EDT

PDB ID	:	4GHF
Title	:	Structure of Y257F variant of Homoprotocatechuate 2,3-Dioxygenase from B.fuscum in complex with 4-Nitrocatechol and dioxygen at 1.67 Ang resolution
Authors	:	Kovaleva, E.G.; Lipscomb, J.D.
Deposited on	:	2012-08-07
Resolution	:	1.67 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

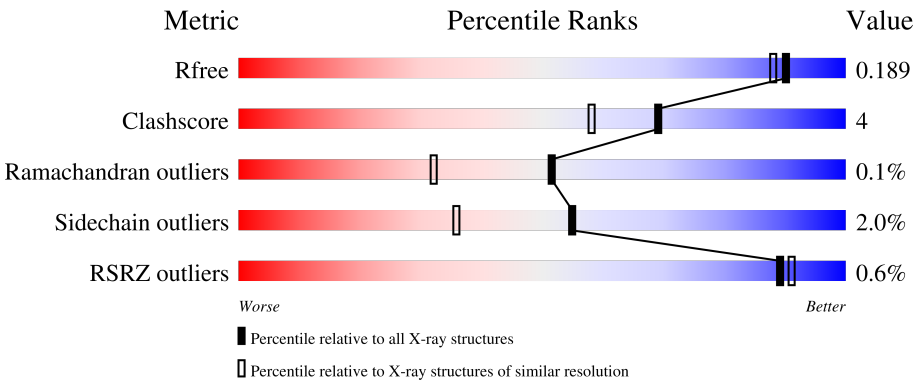
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.67 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	6780 (1.70-1.66)
Clashscore	141614	7310 (1.70-1.66)
Ramachandran outliers	138981	7173 (1.70-1.66)
Sidechain outliers	138945	7172 (1.70-1.66)
RSRZ outliers	127900	6661 (1.70-1.66)

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	365	.% 88% 9% • •
1	B	365	89% 8% • •
1	C	365	2% 88% 8% • •
1	D	365	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	4NC	C	402	-	-	X	-
3	4NC	D	402	-	-	X	-
5	CL	A	404	-	-	X	-
5	CL	C	406	-	-	X	-
5	CL	D	405	-	-	X	-

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 13123 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 Homoprotocatechuate 2,3-dioxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	355	2926	1839	516	564	7	0	7	0
1	B	359	2953	1860	519	566	8	0	7	0
1	C	351	2914	1833	513	559	9	0	9	0
1	D	359	2980	1874	527	571	8	0	10	0

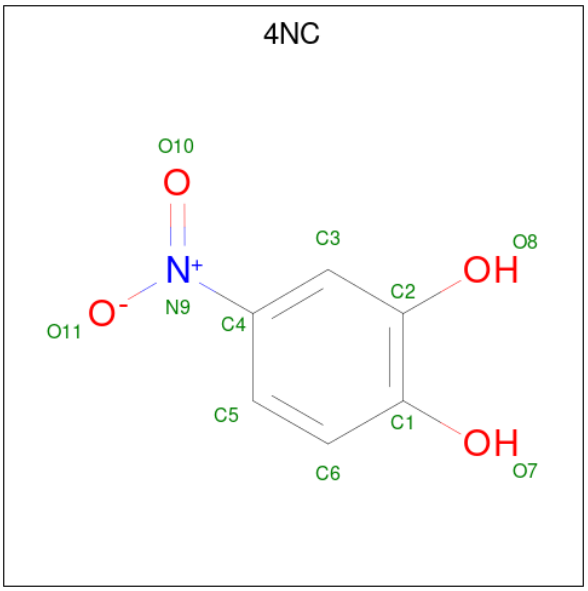
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	257	PHE	TYR	engineered mutation	UNP Q45135
B	257	PHE	TYR	engineered mutation	UNP Q45135
C	257	PHE	TYR	engineered mutation	UNP Q45135
D	257	PHE	TYR	engineered mutation	UNP Q45135

- Molecule 2 is FE (II) ION (three-letter code: FE2) (formula: Fe).

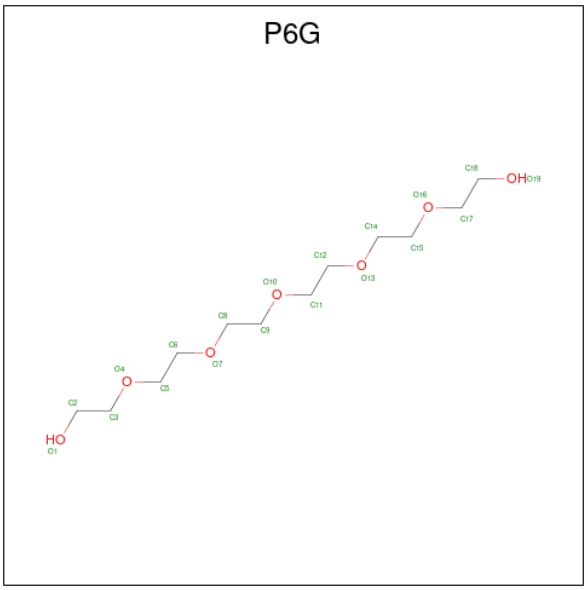
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Fe		
2	A	1	1	1	0	0
2	B	1	1	1	0	0
2	C	1	1	1	0	0
2	D	1	1	1	0	0

- Molecule 3 is 4-NITROCATECHOL (three-letter code: 4NC) (formula: $C_6H_5NO_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	11	6	1	4	0	0
3	C	1	11	6	1	4	0	0
3	D	1	11	6	1	4	0	0

- Molecule 4 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula: $C_{12}H_{26}O_7$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	A	1	19	12	7	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	B	1	19	12	7	0	0
			Total	C	O		
4	B	1	19	12	7	0	0
			Total	C	O		
4	C	1	19	12	7	0	0
			Total	C	O		
4	D	1	19	12	7	0	0

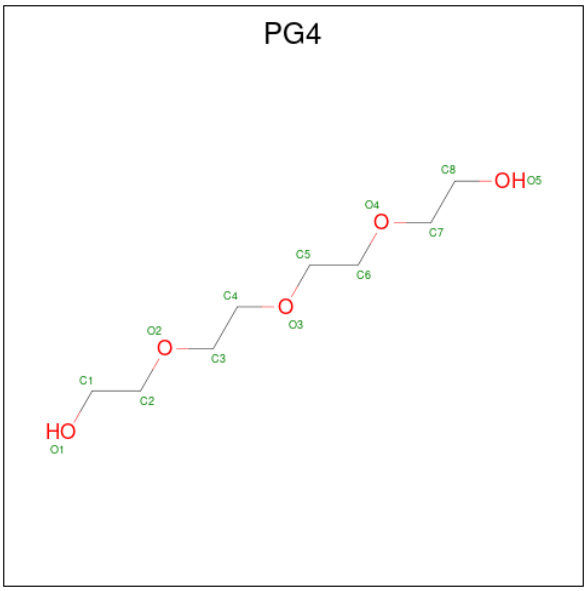
- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
5	A	1	1	1	0	0
			Total	Cl		
5	B	1	1	1	0	0
			Total	Cl		
5	C	1	1	1	0	0
			Total	Cl		
5	D	1	1	1	0	0

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

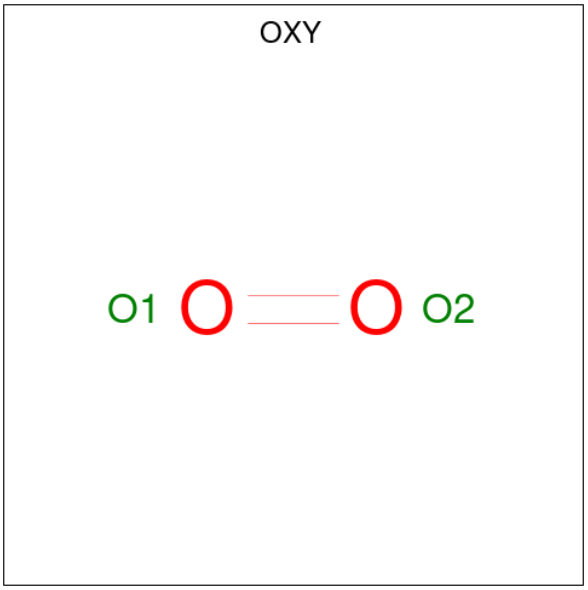
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
6	B	1	1	1	0	0

- Molecule 7 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	1	Total C O 13 8 5	0	0
7	C	1	Total C O 13 8 5	0	0

- Molecule 8 is OXYGEN MOLECULE (three-letter code: OXY) (formula: O_2).



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

- Molecule 9 is water.

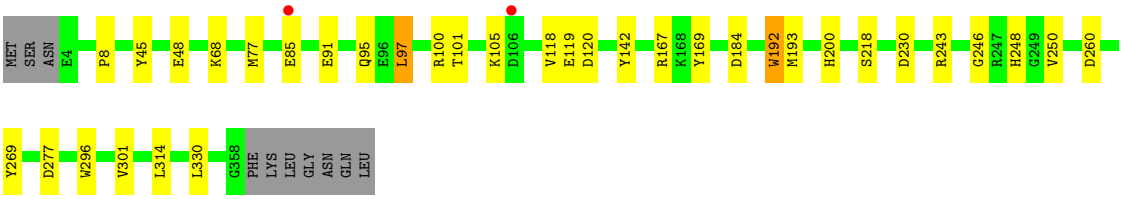
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
9	A	267	269	269	0	2
			Total	O		
9	B	298	300	300	0	2
			Total	O		
9	C	292	296	296	0	4
			Total	O		
9	D	312	318	318	0	6

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.

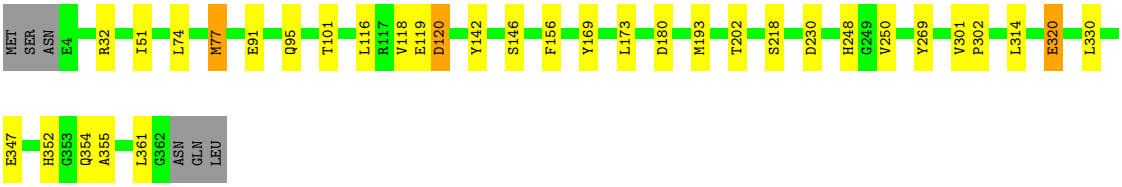
• Molecule 1: Homoprotocatechuate 2,3-dioxygenase

Chain A:



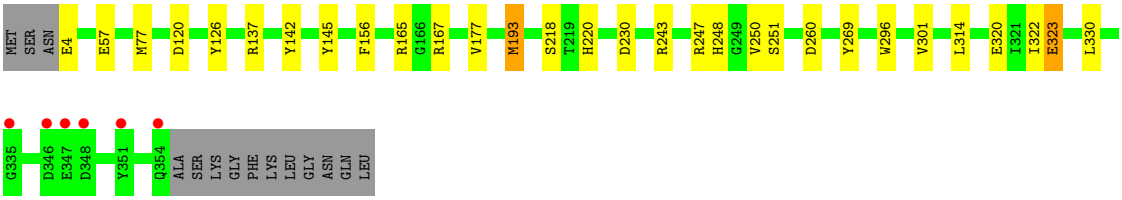
• Molecule 1: Homoprotocatechuate 2,3-dioxygenase

Chain B:



• Molecule 1: Homoprotocatechuate 2,3-dioxygenase

Chain C:



• Molecule 1: Homoprotocatechuate 2,3-dioxygenase

Chain D:

I334
F359
K360
L361
G362
ASN
GLN
LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	110.52Å 150.32Å 96.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	89.04 – 1.67 89.04 – 1.67	Depositor EDS
% Data completeness (in resolution range)	99.5 (89.04-1.67) 99.5 (89.04-1.67)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.97 (at 1.67Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.151 , 0.181 0.160 , 0.189	Depositor DCC
R_{free} test set	9169 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	21.8	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 43.5	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	13123	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 2.97% 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: CA, PG4, 4NC, OXY, CL, P6G, FE2

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	1/3001 (0.0%)	0.85	3/4075 (0.1%)
1	B	0.79	0/3029	0.86	2/4111 (0.0%)
1	C	0.80	1/2989 (0.0%)	0.90	4/4059 (0.1%)
1	D	0.79	0/3056	0.87	4/4146 (0.1%)
All	All	0.78	2/12075 (0.0%)	0.87	13/16391 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	192	TRP	CD2-CE2	6.08	1.48	1.41
1	C	296	TRP	CD2-CE2	5.35	1.47	1.41

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	265	ARG	NE-CZ-NH1	7.30	123.95	120.30
1	C	137[A]	ARG	NE-CZ-NH2	-6.91	116.84	120.30
1	C	137[B]	ARG	NE-CZ-NH2	-6.91	116.84	120.30
1	B	120	ASP	CB-CG-OD1	6.02	123.72	118.30
1	A	167	ARG	NE-CZ-NH2	-5.90	117.35	120.30
1	C	247	ARG	NE-CZ-NH1	5.68	123.14	120.30
1	C	165	ARG	NE-CZ-NH1	5.57	123.09	120.30
1	A	184	ASP	CB-CG-OD1	5.34	123.10	118.30
1	A	277	ASP	CB-CG-OD1	5.29	123.06	118.30
1	D	211	ARG	NE-CZ-NH1	5.23	122.91	120.30
1	D	265	ARG	NE-CZ-NH2	-5.09	117.76	120.30
1	D	180	ASP	CB-CG-OD1	5.07	122.86	118.30
1	B	180	ASP	CB-CG-OD1	5.07	122.86	118.30

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	2926	0	2759	17	1
1	B	2953	0	2794	22	0
1	C	2914	0	2751	22	1
1	D	2980	0	2818	18	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	11	0	3	2	0
3	C	11	0	3	6	0
3	D	11	0	3	5	0
4	A	19	0	26	0	0
4	B	38	0	52	0	0
4	C	19	0	26	0	0
4	D	19	0	26	0	0
5	A	1	0	0	3	0
5	B	1	0	0	0	0
5	C	1	0	0	6	0
5	D	1	0	0	6	0
6	B	1	0	0	0	0
7	B	13	0	18	0	0
7	C	13	0	18	0	0
8	C	2	0	0	0	0
8	D	2	0	0	0	0
9	A	269	0	0	2	0
9	B	300	0	0	4	1
9	C	296	0	0	4	0
9	D	318	0	0	3	1
All	All	13123	0	11297	87	2

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 4.

All (87) 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)
3:C:402:4NC:N9	3:C:402:4NC:O10	1.56	1.39
3:C:402:4NC:O10	5:C:406:CL:CL	0.98	1.35
3:A:402:4NC:O10	5:A:404:CL:CL	0.89	1.25
3:D:402:4NC:N9	5:D:405:CL:CL	2.23	1.06
3:D:402:4NC:O10	5:D:405:CL:CL	0.69	1.05
3:C:402:4NC:N9	5:C:406:CL:CL	2.38	0.93
3:A:402:4NC:N9	5:A:404:CL:CL	2.42	0.89
1:A:243:ARG:NH1	5:A:404:CL:CL	2.48	0.83
1:C:320:GLU:OE2	9:C:521:HOH:O	1.98	0.81
1:C:243:ARG:NH1	5:C:406:CL:CL	2.52	0.76
1:B:301:VAL:HG13	9:B:737:HOH:O	1.88	0.72
1:B:320:GLU:HG2	9:B:661:HOH:O	1.89	0.72
1:A:8:PRO:HG3	1:A:97:LEU:HD13	1.70	0.71
1:D:302:PRO:HG3	1:D:361:LEU:HD22	1.72	0.70
1:B:74:LEU:HD21	1:B:77[A]:MET:HG3	1.74	0.69
1:C:301:VAL:HG22	9:C:676:HOH:O	1.91	0.69
1:C:330:LEU:C	1:C:330:LEU:HD23	2.16	0.66
3:C:402:4NC:H3	5:C:406:CL:CL	2.36	0.63
1:A:91:GLU:O	1:A:95:GLN:HG2	1.99	0.63
1:C:145:TYR:OH	1:C:220:HIS:HD2	1.84	0.60
1:C:177:VAL:HG22	1:C:193[B]:MET:HG2	1.83	0.60
3:C:402:4NC:C3	5:C:406:CL:CL	2.88	0.59
1:A:48[B]:GLU:H	1:A:48[B]:GLU:CD	2.06	0.59
1:A:301:VAL:HG22	9:A:535:HOH:O	2.04	0.56
1:D:243:ARG:NH1	5:D:405:CL:CL	2.73	0.55
1:B:301:VAL:HG12	1:D:359:PHE:CE2	2.41	0.55
1:A:330:LEU:C	1:A:330:LEU:HD23	2.27	0.55
3:D:402:4NC:C3	5:D:405:CL:CL	2.92	0.54
3:D:402:4NC:H3	5:D:405:CL:CL	2.43	0.54
1:C:218[A]:SER:HA	1:C:269:TYR:O	2.07	0.54
1:C:320:GLU:H	1:C:320:GLU:CD	2.11	0.54
1:B:347:GLU:OE2	1:B:352:HIS:NE2	2.34	0.54
1:B:218:SER:HA	1:B:269:TYR:O	2.08	0.53
1:C:77[B]:MET:CE	1:C:156:PHE:CD2	2.92	0.53
1:D:218[B]:SER:HA	1:D:269:TYR:O	2.09	0.52
1:C:218[B]:SER:HA	1:C:269:TYR:O	2.07	0.52
1:B:301:VAL:CG1	9:B:762:HOH:O	2.57	0.52
1:D:101:THR:HG22	1:D:118:VAL:HG12	1.92	0.51
3:D:402:4NC:C4	5:D:405:CL:CL	2.95	0.51
1:A:218[B]:SER:HA	1:A:269:TYR:O	2.12	0.50
1:A:218[A]:SER:HA	1:A:269:TYR:O	2.12	0.50
1:C:167:ARG:HA	1:C:193[B]:MET:HE1	1.93	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:101:THR:HG22	1:A:118:VAL:HG12	1.94	0.48
1:D:248:HIS:HB3	1:D:250:VAL:O	2.13	0.48
1:D:325:THR:O	9:D:780[A]:HOH:O	2.20	0.48
1:D:218[A]:SER:HA	1:D:269:TYR:O	2.12	0.48
1:B:301:VAL:HG11	9:B:762:HOH:O	2.14	0.48
1:B:91:GLU:HG2	1:B:95:GLN:NE2	2.29	0.48
1:B:302:PRO:CG	1:B:361:LEU:HD23	2.43	0.48
1:A:248:HIS:HB3	1:A:250:VAL:O	2.13	0.47
1:A:100:ARG:HG2	1:A:314:LEU:HD12	1.96	0.47
1:B:74:LEU:HD21	1:B:77[A]:MET:CG	2.44	0.47
1:D:246:GLY:HA2	1:D:296:TRP:CZ3	2.50	0.47
1:C:77[A]:MET:HE2	1:C:126:TYR:CE1	2.50	0.47
1:B:142:TYR:CG	1:C:230:ASP:HB3	2.50	0.47
1:C:77[B]:MET:HE2	1:C:156:PHE:CD2	2.49	0.47
1:A:230:ASP:HB3	1:D:142:TYR:CG	2.51	0.46
1:B:101:THR:HG22	1:B:118:VAL:HG12	1.96	0.46
1:D:176:ARG:HD2	9:D:731:HOH:O	2.14	0.46
3:C:402:4NC:C4	5:C:406:CL:CL	3.00	0.46
1:A:68:LYS:NZ	9:A:612:HOH:O	2.33	0.46
1:C:248:HIS:HB3	1:C:250:VAL:O	2.16	0.45
1:B:248:HIS:HB3	1:B:250:VAL:O	2.16	0.45
1:B:330:LEU:C	1:B:330:LEU:HD23	2.37	0.45
1:B:354[A]:GLN:HG3	1:B:355:ALA:O	2.16	0.45
1:A:142:TYR:CG	1:D:230:ASP:HB3	2.52	0.45
1:C:330:LEU:HD23	1:C:330:LEU:O	2.17	0.44
1:A:192:TRP:HD1	1:A:200:HIS:CD2	2.36	0.44
1:B:101:THR:HB	1:B:116:LEU:HD11	2.00	0.43
1:A:119:GLU:HG3	1:A:314:LEU:HG	2.01	0.43
1:C:330:LEU:C	1:C:330:LEU:CD2	2.86	0.43
1:D:151:VAL:HG11	1:D:220:HIS:CE1	2.53	0.43
1:C:322:ILE:HD13	9:C:708:HOH:O	2.20	0.42
1:D:167[B]:ARG:HD3	9:D:616:HOH:O	2.19	0.42
1:A:246:GLY:HA2	1:A:296:TRP:CZ3	2.55	0.42
1:C:4:GLU:N	9:C:789:HOH:O	2.52	0.42
1:B:302:PRO:HG2	1:B:361:LEU:HD23	2.00	0.42
1:D:334:ILE:HD12	1:D:361:LEU:HD21	2.02	0.42
1:B:119:GLU:HG3	1:B:314:LEU:HG	2.02	0.42
1:B:156:PHE:O	1:B:202:THR:HA	2.20	0.41
1:C:77[B]:MET:HE3	1:C:156:PHE:CD2	2.55	0.41
1:B:32:ARG:HB2	1:B:51:ILE:CD1	2.50	0.41
1:D:50:GLN:HA	1:D:66:LEU:O	2.21	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:119:GLU:HG3	1:D:314:LEU:HG	2.02	0.40
1:B:230:ASP:HB3	1:C:142:TYR:CG	2.56	0.40
1:D:101:THR:HB	1:D:116:LEU:HD11	2.03	0.40
1:C:57:GLU:OE2	1:C:220:HIS:HE1	2.04	0.40

All (2) 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)
1:A:45:TYR:OH	9:B:634:HOH:O[4_555]	1.81	0.39
1:C:323[A]:GLU:OE1	9:D:811:HOH:O[4_456]	2.12	0.08

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	360/365 (99%)	350 (97%)	10 (3%)	0	100	100
1	B	364/365 (100%)	357 (98%)	7 (2%)	0	100	100
1	C	358/365 (98%)	351 (98%)	6 (2%)	1 (0%)	41	23
1	D	367/365 (100%)	358 (98%)	9 (2%)	0	100	100
All	All	1449/1460 (99%)	1416 (98%)	32 (2%)	1 (0%)	51	32

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	251	SER

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	311/313 (99%)	303 (97%)	8 (3%)	46	25
1	B	314/313 (100%)	306 (98%)	8 (2%)	47	26
1	C	311/313 (99%)	304 (98%)	7 (2%)	50	30
1	D	317/313 (101%)	312 (98%)	5 (2%)	62	46
All	All	1253/1252 (100%)	1225 (98%)	28 (2%)	55	32

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	77	MET
1	A	85	GLU
1	A	97	LEU
1	A	105	LYS
1	A	120	ASP
1	A	169	TYR
1	A	193	MET
1	A	260	ASP
1	B	77[A]	MET
1	B	77[B]	MET
1	B	120	ASP
1	B	146	SER
1	B	169	TYR
1	B	173	LEU
1	B	193	MET
1	B	320	GLU
1	C	120	ASP
1	C	193[A]	MET
1	C	193[B]	MET
1	C	260	ASP
1	C	314	LEU
1	C	323[A]	GLU
1	C	323[B]	GLU
1	D	105	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	120	ASP
1	D	176	ARG
1	D	193	MET
1	D	260	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	B	95	GLN
1	C	95	GLN
1	C	220	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)

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 21 ligands modelled in this entry, 9 are monoatomic - leaving 12 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	4NC	D	402	2	10,11,11	6.37	4 (40%)	13,15,15	2.24	2 (15%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	P6G	B	405	-	18,18,18	0.57	0	17,17,17	0.62	0
3	4NC	C	402	2	10,11,11	6.51	3 (30%)	13,15,15	0.79	0
4	P6G	D	403	-	18,18,18	0.64	0	17,17,17	0.35	0
8	OXY	D	404	2	1,1,1	0.16	0	-		
8	OXY	C	405	2	1,1,1	0.17	0	-		
3	4NC	A	402	2	10,11,11	6.35	4 (40%)	13,15,15	1.33	2 (15%)
4	P6G	C	404	-	18,18,18	0.49	0	17,17,17	0.58	0
7	PG4	B	404	-	12,12,12	0.53	0	11,11,11	0.41	0
4	P6G	A	403	-	18,18,18	0.57	0	17,17,17	0.46	0
7	PG4	C	403	-	12,12,12	0.64	0	11,11,11	0.33	0
4	P6G	B	406	-	18,18,18	0.63	0	17,17,17	0.52	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	4NC	D	402	2	-	1/2/4/4	0/1/1/1
4	P6G	B	405	-	-	3/16/16/16	-
3	4NC	C	402	2	-	2/2/4/4	0/1/1/1
4	P6G	D	403	-	-	1/16/16/16	-
3	4NC	A	402	2	-	2/2/4/4	0/1/1/1
4	P6G	C	404	-	-	3/16/16/16	-
7	PG4	B	404	-	-	5/10/10/10	-
4	P6G	A	403	-	-	4/16/16/16	-
7	PG4	C	403	-	-	1/10/10/10	-
4	P6G	B	406	-	-	2/16/16/16	-

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	402	4NC	O10-N9	19.89	1.56	1.22
3	D	402	4NC	O10-N9	19.20	1.55	1.22
3	A	402	4NC	O10-N9	18.89	1.54	1.22
3	D	402	4NC	C2-C1	4.57	1.47	1.40
3	A	402	4NC	C2-C1	4.55	1.47	1.40
3	A	402	4NC	C4-N9	-4.11	1.35	1.45
3	C	402	4NC	C2-C1	3.81	1.46	1.40
3	C	402	4NC	C4-N9	-3.25	1.37	1.45

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	402	4NC	C4-N9	-2.36	1.39	1.45
3	A	402	4NC	O8-C2	-2.15	1.32	1.36
3	D	402	4NC	C6-C5	2.10	1.42	1.38

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	402	4NC	O10-N9-C4	-6.58	109.48	118.80
3	D	402	4NC	C5-C6-C1	-3.79	116.61	120.50
3	A	402	4NC	C5-C6-C1	-2.83	117.59	120.50
3	A	402	4NC	O10-N9-C4	-2.14	115.76	118.80

There are no chirality outliers.

All (24) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	402	4NC	C5-C4-N9-O10
4	B	406	P6G	O16-C17-C18-O19
7	B	404	PG4	O1-C1-C2-O2
7	B	404	PG4	O4-C7-C8-O5
3	C	402	4NC	C5-C4-N9-O10
3	D	402	4NC	C5-C4-N9-O10
4	B	405	P6G	O1-C2-C3-O4
4	A	403	P6G	O4-C5-C6-O7
4	C	404	P6G	O1-C2-C3-O4
7	B	404	PG4	C4-C3-O2-C2
4	B	405	P6G	C6-C5-O4-C3
7	B	404	PG4	C1-C2-O2-C3
4	A	403	P6G	C5-C6-O7-C8
4	C	404	P6G	C6-C5-O4-C3
3	A	402	4NC	C3-C4-N9-O10
4	D	403	P6G	C5-C6-O7-C8
4	A	403	P6G	C9-C8-O7-C6
4	A	403	P6G	C2-C3-O4-C5
4	C	404	P6G	C5-C6-O7-C8
7	B	404	PG4	C5-C6-O4-C7
4	B	406	P6G	C5-C6-O7-C8
7	C	403	PG4	C8-C7-O4-C6
4	B	405	P6G	C9-C8-O7-C6
3	C	402	4NC	C3-C4-N9-O10

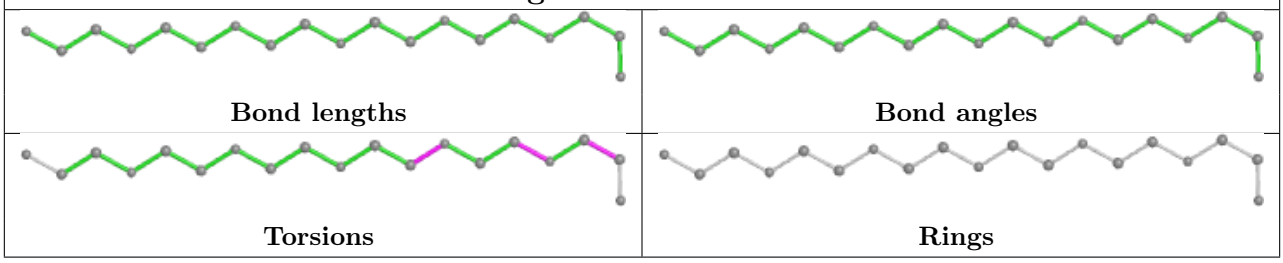
There are no ring outliers.

3 monomers are involved in 13 short contacts:

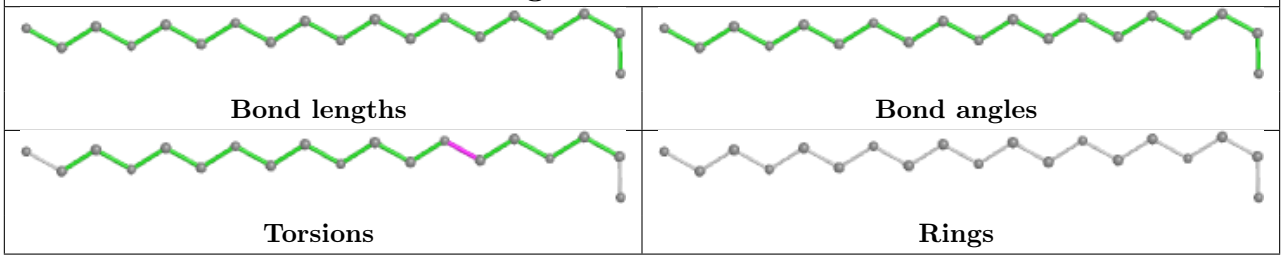
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	402	4NC	5	0
3	C	402	4NC	6	0
3	A	402	4NC	2	0

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.

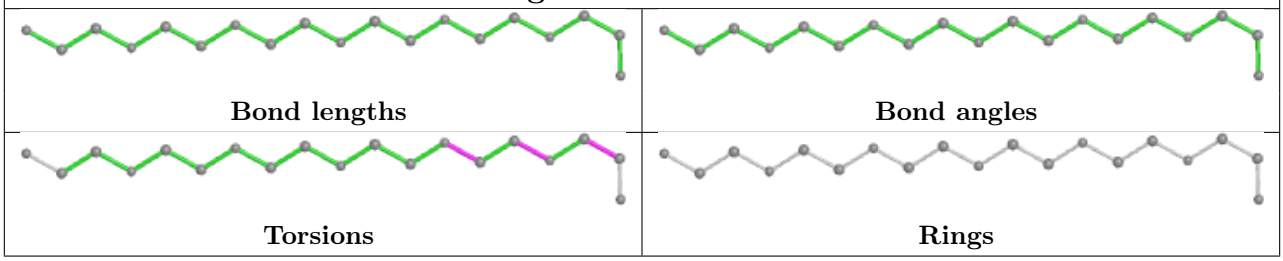
Ligand P6G B 405

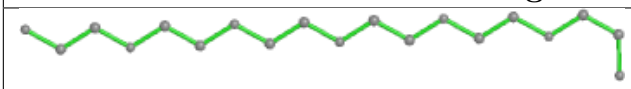
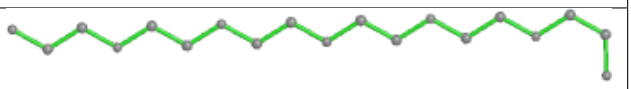
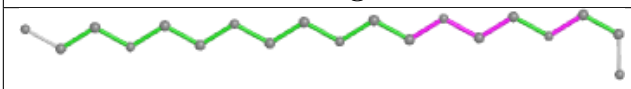
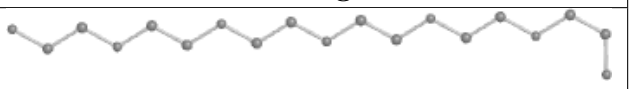


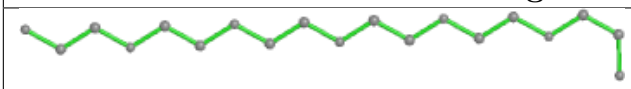
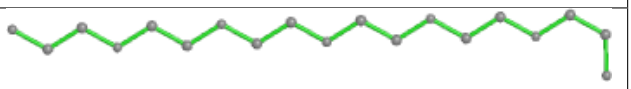
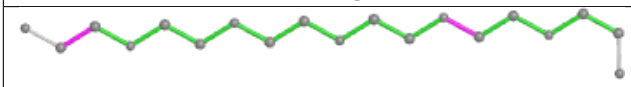
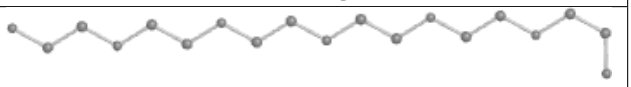
Ligand P6G D 403



Ligand P6G C 404



Ligand P6G A 403	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand P6G B 406	
 Bond lengths	 Bond angles
 Torsions	 Rings

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	355/365 (97%)	-0.29	2 (0%) 89 91	16, 28, 47, 58	0
1	B	359/365 (98%)	-0.35	0 100 100	16, 25, 38, 57	0
1	C	351/365 (96%)	-0.37	6 (1%) 70 74	15, 23, 40, 73	0
1	D	359/365 (98%)	-0.40	1 (0%) 94 94	15, 24, 40, 59	0
All	All	1424/1460 (97%)	-0.35	9 (0%) 89 91	15, 25, 43, 73	0

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	354	GLN	3.0
1	A	85	GLU	2.5
1	C	351	TYR	2.4
1	D	176	ARG	2.3
1	C	348	ASP	2.3
1	C	335	GLY	2.2
1	C	346	ASP	2.1
1	A	106	ASP	2.1
1	C	347	GLU	2.1

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)

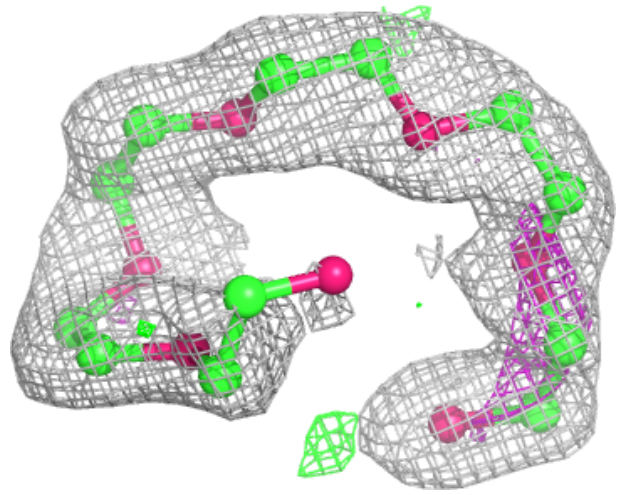
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

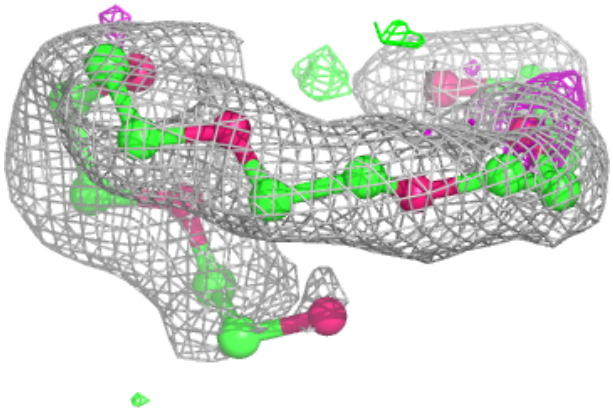
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	PG4	C	403	13/13	0.82	0.18	51,56,63,63	0
7	PG4	B	404	13/13	0.84	0.16	46,51,56,60	0
4	P6G	A	403	19/19	0.87	0.12	29,40,66,68	0
4	P6G	B	406	19/19	0.89	0.12	35,41,49,49	0
3	4NC	D	402	11/11	0.90	0.12	17,20,26,26	11
4	P6G	D	403	19/19	0.91	0.09	26,33,51,60	0
4	P6G	C	404	19/19	0.93	0.09	24,31,51,54	0
3	4NC	C	402	11/11	0.94	0.12	20,21,28,31	0
3	4NC	A	402	11/11	0.94	0.13	20,23,30,33	11
4	P6G	B	405	19/19	0.95	0.08	27,32,52,53	0
5	CL	C	406	1/1	0.97	0.10	17,17,17,17	1
5	CL	D	405	1/1	0.98	0.06	19,19,19,19	1
6	CA	B	403	1/1	0.99	0.06	21,21,21,21	0
5	CL	A	404	1/1	0.99	0.12	19,19,19,19	1
5	CL	B	402	1/1	0.99	0.05	26,26,26,26	0
8	OXY	C	405	2/2	0.99	0.10	20,20,20,22	2
8	OXY	D	404	2/2	0.99	0.08	26,26,26,26	2
2	FE2	B	401	1/1	1.00	0.07	21,21,21,21	0
2	FE2	C	401	1/1	1.00	0.05	21,21,21,21	0
2	FE2	D	401	1/1	1.00	0.05	20,20,20,20	0
2	FE2	A	401	1/1	1.00	0.08	23,23,23,23	0

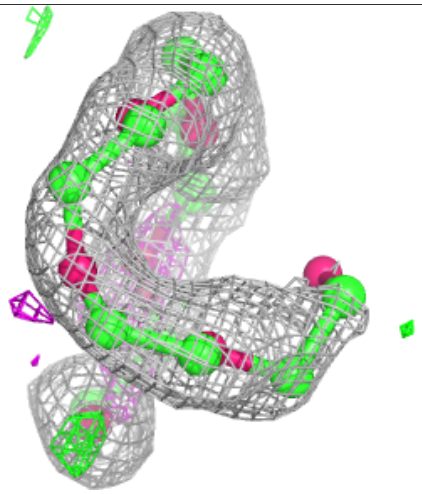
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around P6G A 403:

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

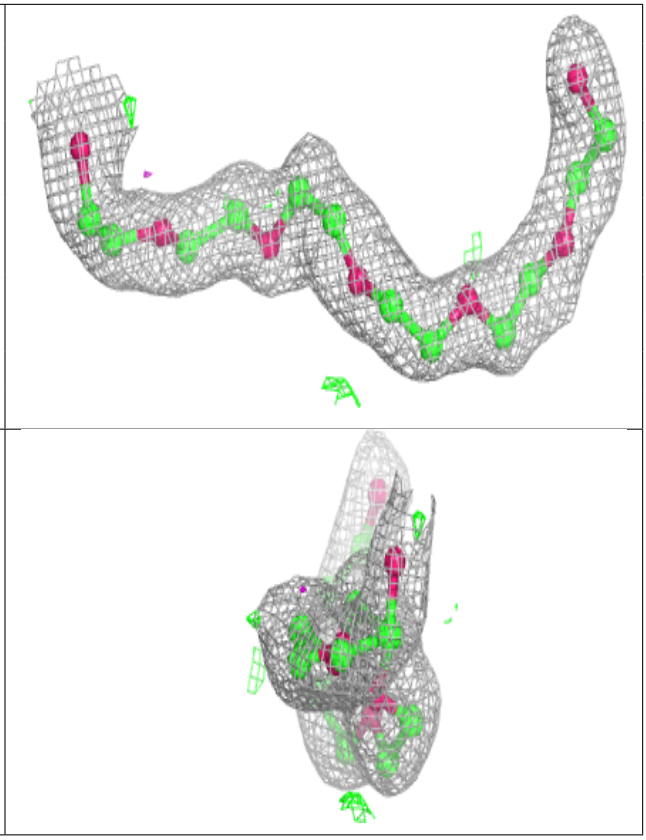






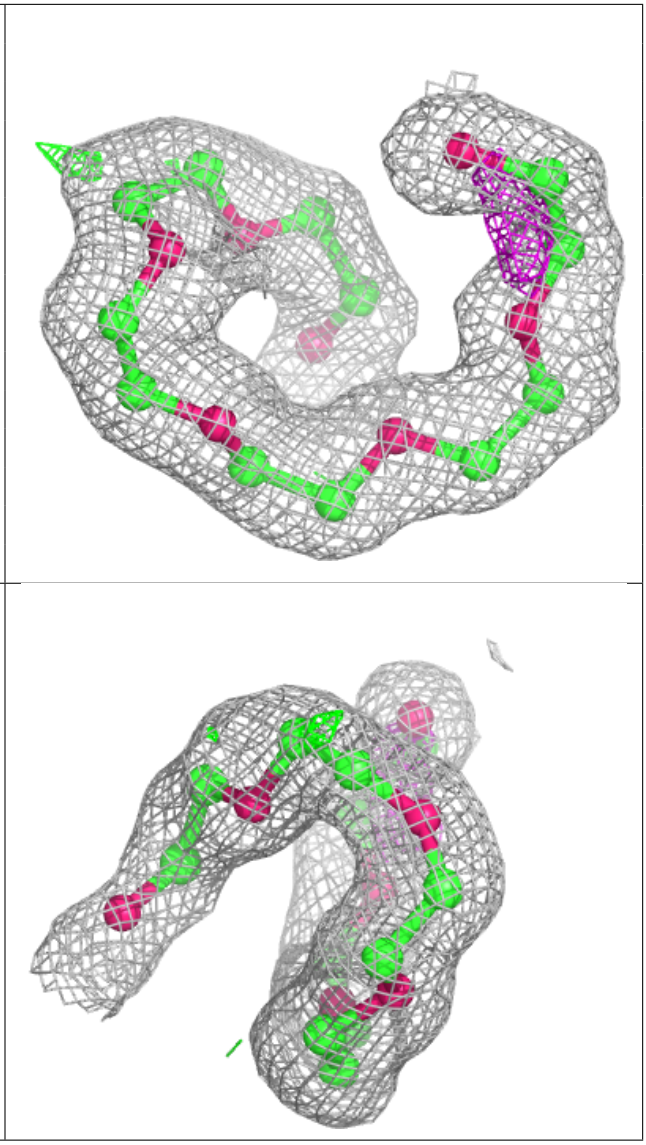
Electron density around P6G B 406:

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



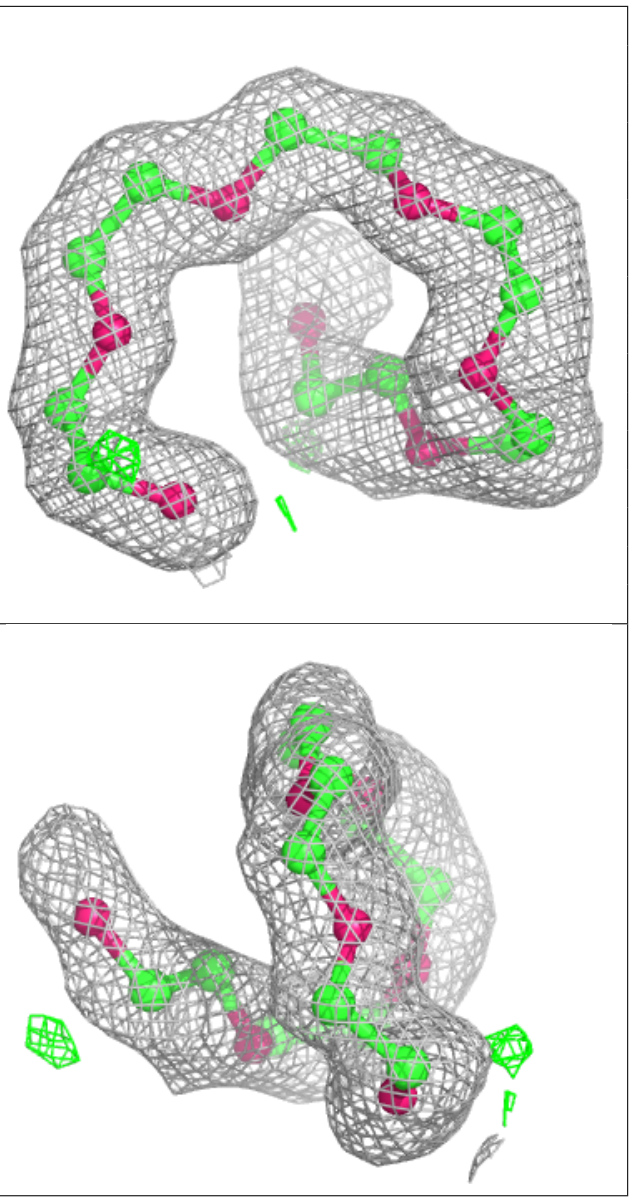
Electron density around P6G D 403:

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



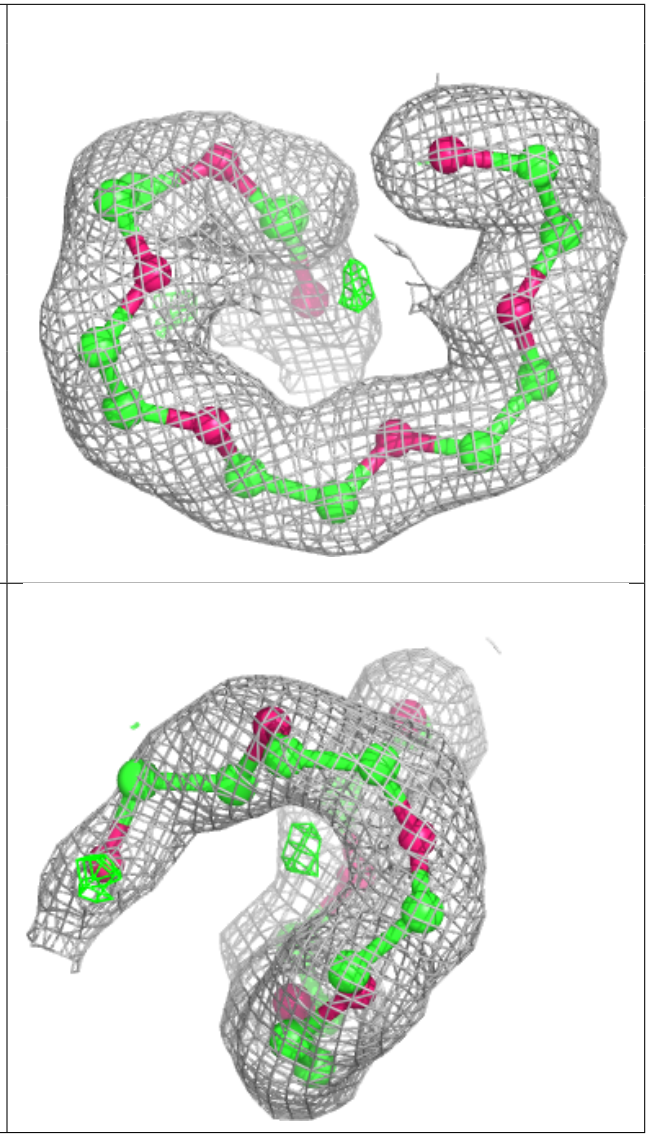
Electron density around P6G C 404:

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



Electron density around P6G B 405:

$2mF_o$-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.