wwPDB X-ray Structure Validation Summary Report (i)

Oct 20, 2024 – 04:09 PM EDT

PDB ID	:	1G8K
Title	:	CRYSTAL STRUCTURE ANALYSIS OF ARSENITE OXIDASE FROM ALCALIGENES FAECALIS
Authors	:	Ellis, P.J.; Conrads, T.; Hille, R.; Kuhn, P.
Deposited on	:	2000-11-17
Resolution	:	1.64 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

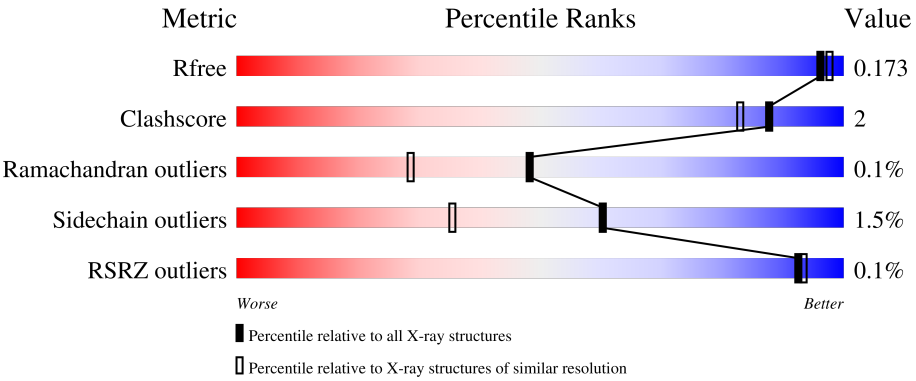
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 1.64 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}	164625	1015 (1.64-1.64)
Clashscore	180529	1093 (1.64-1.64)
Ramachandran outliers	177936	1077 (1.64-1.64)
Sidechain outliers	177891	1077 (1.64-1.64)
RSRZ outliers	164620	1015 (1.64-1.64)

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	825	95% 5%
1	C	825	94% 5%
1	E	825	93% 6%
1	G	825	94% 6%
2	B	133	98%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	133	97%
2	F	133	97%
2	H	133	% 97%

2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 34251 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 ARSENITE OXIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	822	Total 6455	C 4060	N 1136	O 1219	S 40	67	0	0
1	C	822	Total 6455	C 4060	N 1136	O 1219	S 40	44	0	0
1	E	822	Total 6455	C 4060	N 1136	O 1219	S 40	37	0	0
1	G	822	Total 6455	C 4060	N 1136	O 1219	S 40	61	0	0

- Molecule 2 is a protein called ARSENITE OXIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	133	Total 965	C 604	N 158	O 195	S 8	0	0	0
2	D	133	Total 965	C 604	N 158	O 195	S 8	0	0	0
2	F	133	Total 965	C 604	N 158	O 195	S 8	0	0	0
2	H	133	Total 965	C 604	N 158	O 195	S 8	0	0	0

- Molecule 3 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

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

Continued on next page...

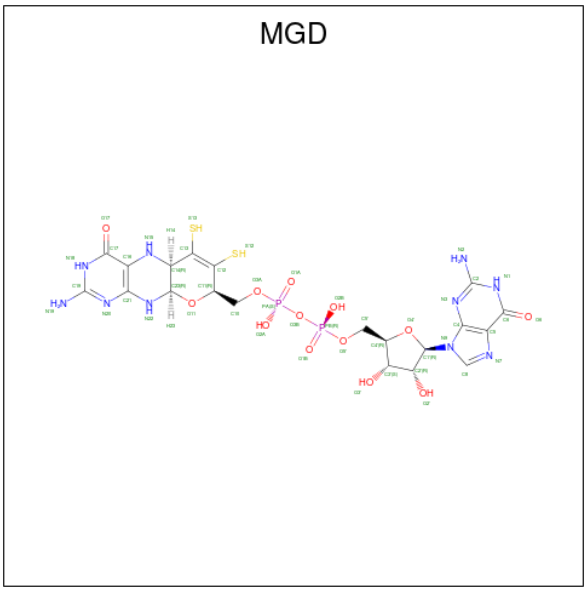
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	E	2	Total Hg 2 2	0	0
3	F	1	Total Hg 1 1	0	0
3	G	2	Total Hg 2 2	0	0
3	H	1	Total Hg 1 1	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Ca 2 2	0	0
4	C	2	Total Ca 2 2	0	0
4	E	2	Total Ca 2 2	0	0
4	G	2	Total Ca 2 2	0	0

- Molecule 5 is 2-AMINO-5,6-DIMERCAPTO-7-METHYL-3,7,8A,9-TETRAHYDRO-8-OXA-1,3,9,10-TETRAAZA-ANTHRACEN-4-ONE GUANOSINE DINUCLEOTIDE (three-letter code: MGD) (formula: $C_{20}H_{26}N_{10}O_{13}P_2S_2$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
5	A	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
5	C	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
5	C	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
5	E	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
5	E	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
5	G	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		
5	G	1	Total	C	N	O	P	S	0	0
			47	20	10	13	2	2		

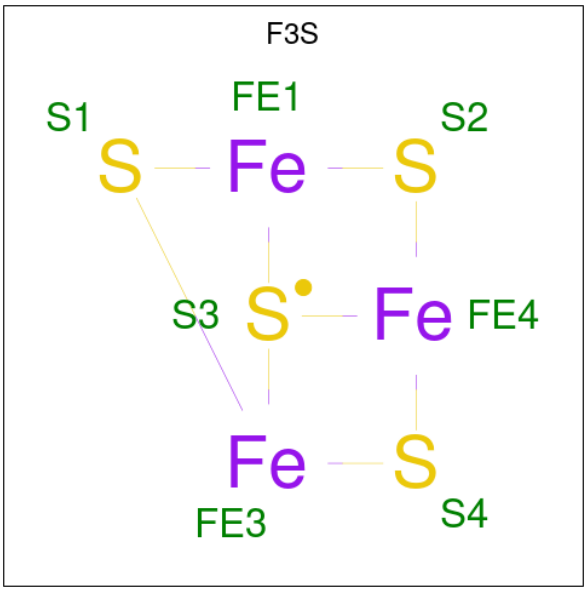
- Molecule 6 is OXYGEN ATOM (three-letter code: O) (formula: O).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	O	0	0
			1	1		
6	C	1	Total	O	0	0
			1	1		
6	E	1	Total	O	0	0
			1	1		
6	G	1	Total	O	0	0
			1	1		

- Molecule 7 is MOLYBDENUM(IV) ION (three-letter code: 4MO) (formula: Mo).

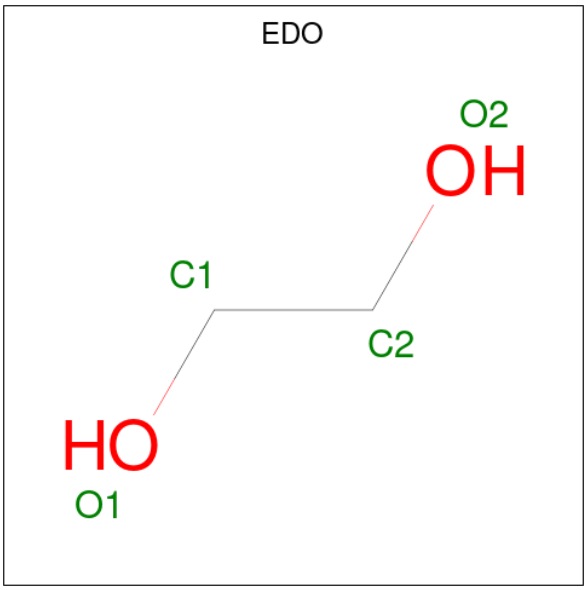
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Mo	0	0
			1	1		
7	C	1	Total	Mo	0	0
			1	1		
7	E	1	Total	Mo	0	0
			1	1		
7	G	1	Total	Mo	0	0
			1	1		

- Molecule 8 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe_3S_4).



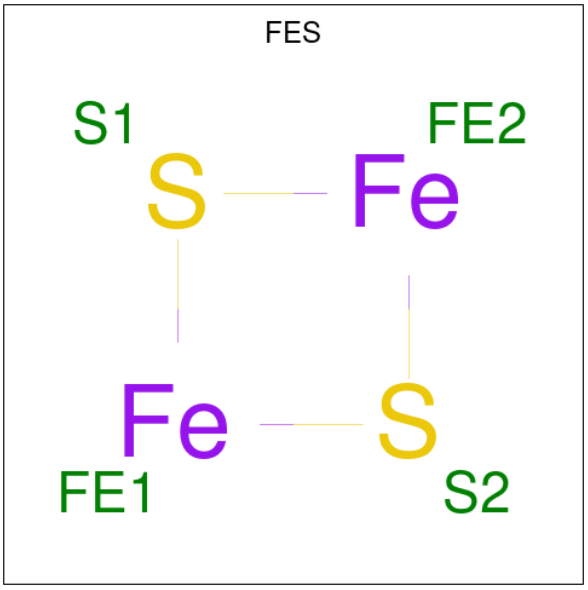
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total Fe S 7 3 4	0	0
8	C	1	Total Fe S 7 3 4	0	0
8	E	1	Total Fe S 7 3 4	0	0
8	G	1	Total Fe S 7 3 4	0	0

- Molecule 9 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 10 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe_2S_2).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	B	1	Total Fe S 4 2 2	0	0
10	D	1	Total Fe S 4 2 2	0	0
10	F	1	Total Fe S 4 2 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	H	1	Total Fe S 4 2 2	0	0

- Molecule 11 is water.

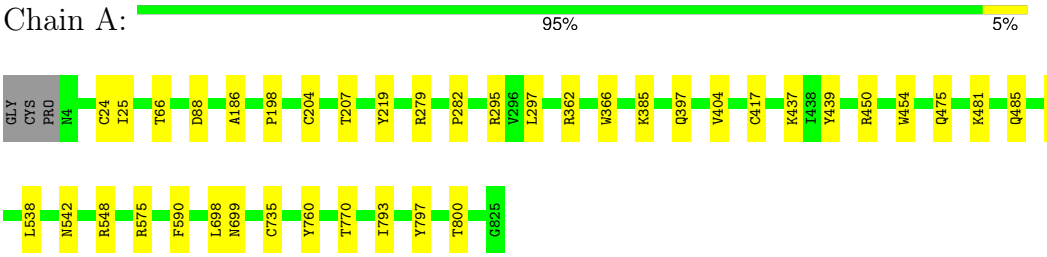
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	804	Total O 804 804	0	0
11	B	165	Total O 165 165	0	0
11	C	908	Total O 908 908	0	0
11	D	157	Total O 157 157	0	0
11	E	878	Total O 878 878	0	0
11	F	162	Total O 162 162	0	0
11	G	849	Total O 849 849	0	0
11	H	164	Total O 164 164	0	0

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.

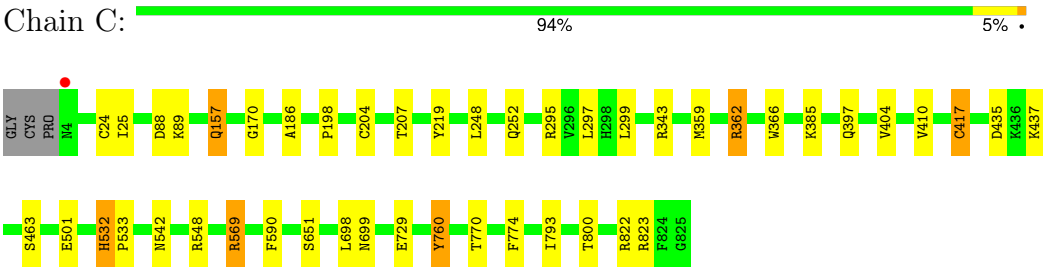
• Molecule 1: ARSENITE OXIDASE

Chain A:



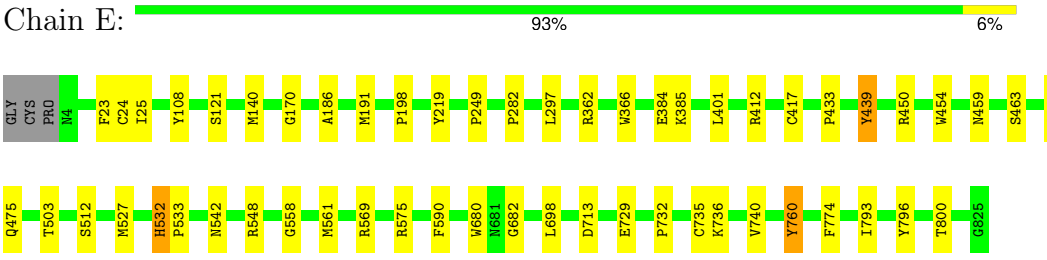
• Molecule 1: ARSENITE OXIDASE

Chain C:



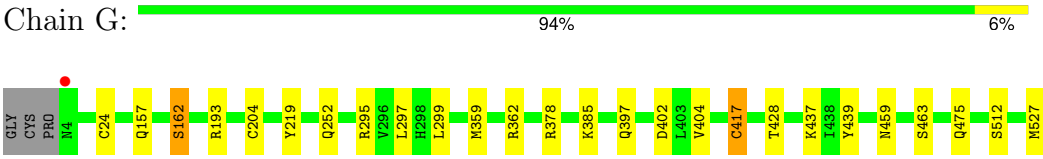
• Molecule 1: ARSENITE OXIDASE

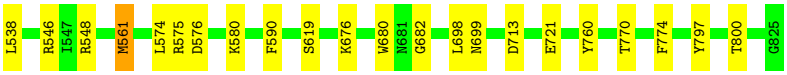
Chain E:



• Molecule 1: ARSENITE OXIDASE

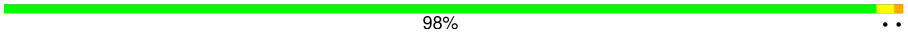
Chain G:





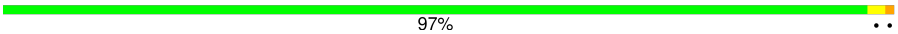
• Molecule 2: ARSENITE OXIDASE

Chain B:



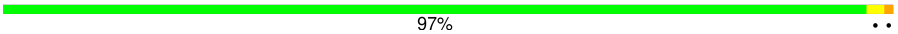
• Molecule 2: ARSENITE OXIDASE

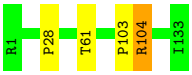
Chain D:



• Molecule 2: ARSENITE OXIDASE

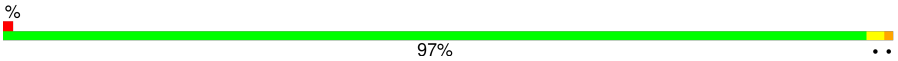
Chain F:





• Molecule 2: ARSENITE OXIDASE

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	90.74Å 109.52Å 117.64Å 97.71° 90.00° 96.43°	Depositor
Resolution (Å)	17.80 – 1.64 17.80 – 1.64	Depositor EDS
% Data completeness (in resolution range)	(Not available) (17.80-1.64) 93.5 (17.80-1.64)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.05 (at 1.65Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.154 , 0.179 0.149 , 0.173	Depositor DCC
R_{free} test set	25621 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	13.2	Xtriage
Anisotropy	0.259	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 52.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	34251	wwPDB-VP
Average B, all atoms (Å^2)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.15% 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: MGD, CA, 4MO, O, EDO, HG, F3S, FES

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	0.91	1/6616 (0.0%)	0.90	4/8969 (0.0%)
1	C	0.95	0/6616	0.95	8/8969 (0.1%)
1	E	1.00	5/6616 (0.1%)	0.97	5/8969 (0.1%)
1	G	0.96	0/6616	0.95	8/8969 (0.1%)
2	B	0.98	0/986	0.95	0/1349
2	D	0.93	0/986	0.94	0/1349
2	F	0.93	0/986	0.93	0/1349
2	H	0.98	0/986	0.94	0/1349
All	All	0.96	6/30408 (0.0%)	0.94	25/41272 (0.1%)

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	2
1	C	0	1
1	E	0	2
1	G	0	2
All	All	0	7

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	735	CYS	CB-SG	-7.09	1.70	1.82
1	E	140	MET	CG-SD	5.96	1.96	1.81
1	E	108	TYR	CE1-CZ	5.50	1.45	1.38
1	E	191	MET	SD-CE	-5.48	1.47	1.77
1	E	23	PHE	CD1-CE1	5.41	1.50	1.39

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	295	ARG	NE-CZ-NH1	7.46	124.03	120.30
1	G	295	ARG	NE-CZ-NH2	-6.96	116.82	120.30
1	C	343	ARG	NE-CZ-NH1	6.60	123.60	120.30
1	C	295	ARG	NE-CZ-NH2	-6.54	117.03	120.30
1	E	713	ASP	CB-CG-OD2	-6.30	112.63	118.30

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	219	TYR	Sidechain
1	A	797	TYR	Sidechain
1	C	219	TYR	Sidechain
1	E	219	TYR	Sidechain
1	E	439	TYR	Sidechain

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	6455	0	6220	20	0
1	C	6455	0	6221	23	0
1	E	6455	0	6220	25	0
1	G	6455	0	6221	24	0
2	B	965	0	934	3	0
2	D	965	0	933	3	0
2	F	965	0	933	2	0
2	H	965	0	933	6	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
3	C	2	0	0	0	0
3	D	1	0	0	0	0
3	E	2	0	0	0	0
3	F	1	0	0	0	0
3	G	2	0	0	1	0
3	H	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	2	0	0	0	0
4	C	2	0	0	0	0
4	E	2	0	0	0	0
4	G	2	0	0	0	0
5	A	94	0	44	2	0
5	C	94	0	44	2	0
5	E	94	0	44	2	0
5	G	94	0	44	1	0
6	A	1	0	0	1	0
6	C	1	0	0	1	0
6	E	1	0	0	1	0
6	G	1	0	0	1	0
7	A	1	0	0	1	0
7	C	1	0	0	1	0
7	E	1	0	0	1	0
7	G	1	0	0	1	0
8	A	7	0	0	0	0
8	C	7	0	0	0	0
8	E	7	0	0	0	0
8	G	7	0	0	0	0
9	A	8	0	12	0	0
9	C	8	0	12	0	0
9	E	12	0	18	0	0
9	G	8	0	12	0	0
10	B	4	0	0	0	0
10	D	4	0	0	0	0
10	F	4	0	0	0	0
10	H	4	0	0	0	0
11	A	804	0	0	3	0
11	B	165	0	0	3	0
11	C	908	0	0	5	0
11	D	157	0	0	1	0
11	E	878	0	0	2	0
11	F	162	0	0	0	0
11	G	849	0	0	8	0
11	H	164	0	0	4	0
All	All	34251	0	28845	110	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 110 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:157:GLN:HG2	11:C:7662:HOH:O	1.72	0.89
1:E:561:MET:HE1	1:E:569:ARG:HH11	1.46	0.80
6:E:5203:O:O	7:E:5204:4MO:MO	1.53	0.77
6:C:5103:O:O	7:C:5104:4MO:MO	1.59	0.73
1:A:475:GLN:HG3	11:A:7620:HOH:O	1.89	0.73

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	820/825 (99%)	790 (96%)	29 (4%)	1 (0%)	48	29
1	C	820/825 (99%)	792 (97%)	27 (3%)	1 (0%)	48	29
1	E	820/825 (99%)	790 (96%)	29 (4%)	1 (0%)	48	29
1	G	820/825 (99%)	792 (97%)	28 (3%)	0	100	100
2	B	131/133 (98%)	125 (95%)	6 (5%)	0	100	100
2	D	131/133 (98%)	130 (99%)	1 (1%)	0	100	100
2	F	131/133 (98%)	127 (97%)	4 (3%)	0	100	100
2	H	131/133 (98%)	127 (97%)	4 (3%)	0	100	100
All	All	3804/3832 (99%)	3673 (97%)	128 (3%)	3 (0%)	48	29

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	793	ILE
1	E	793	ILE
1	A	793	ILE

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	674/676 (100%)	665 (99%)	9 (1%)	65	40
1	C	674/676 (100%)	661 (98%)	13 (2%)	52	24
1	E	674/676 (100%)	663 (98%)	11 (2%)	58	32
1	G	674/676 (100%)	666 (99%)	8 (1%)	67	45
2	B	105/105 (100%)	104 (99%)	1 (1%)	73	54
2	D	105/105 (100%)	104 (99%)	1 (1%)	73	54
2	F	105/105 (100%)	103 (98%)	2 (2%)	52	24
2	H	105/105 (100%)	104 (99%)	1 (1%)	73	54
All	All	3116/3124 (100%)	3070 (98%)	46 (2%)	60	35

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	439	TYR
2	F	28	PRO
1	E	454	TRP
1	E	760	TYR
1	G	24	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	579	GLN
1	G	234	ASN
1	E	689	GLN
1	G	475	GLN
1	C	223	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 53 ligands modelled in this entry, 28 are monoatomic - leaving 25 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	$\#\|Z\| > 2$
5	MGD	E	5201	7	43,52,52	1.49	7 (16%)	43,81,81	1.47	4 (9%)
8	F3S	E	5205	1	0,9,9	-	-	-		
9	EDO	A	7005	-	3,3,3	0.43	0	2,2,2	0.18	0
9	EDO	C	7002	-	3,3,3	0.99	0	2,2,2	0.34	0
5	MGD	G	5301	7	43,52,52	1.65	9 (20%)	43,81,81	1.75	6 (13%)
9	EDO	G	7004	-	3,3,3	0.94	0	2,2,2	0.23	0
5	MGD	C	5101	7	43,52,52	1.59	10 (23%)	43,81,81	1.55	6 (13%)
5	MGD	A	5001	7	43,52,52	1.43	10 (23%)	43,81,81	1.67	7 (16%)
8	F3S	G	5305	1	0,9,9	-	-	-		
5	MGD	G	5302	7	43,52,52	1.34	8 (18%)	43,81,81	1.81	8 (18%)
9	EDO	E	7003	-	3,3,3	0.67	0	2,2,2	0.17	0
10	FES	H	5306	2	0,4,4	-	-	-		
8	F3S	A	5005	1	0,9,9	-	-	-		
10	FES	D	5106	2	0,4,4	-	-	-		
9	EDO	C	7006	-	3,3,3	0.59	0	2,2,2	0.14	0
9	EDO	A	7001	-	3,3,3	0.96	0	2,2,2	0.23	0
9	EDO	E	7007	-	3,3,3	0.50	0	2,2,2	0.30	0
5	MGD	E	5202	7	43,52,52	1.61	13 (30%)	43,81,81	1.79	9 (20%)
10	FES	B	5006	2	0,4,4	-	-	-		
10	FES	F	5206	2	0,4,4	-	-	-		
5	MGD	A	5002	7	43,52,52	1.54	8 (18%)	43,81,81	1.82	7 (16%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
9	EDO	E	7009	-	3,3,3	0.79	0	2,2,2	0.15	0
8	F3S	C	5105	1	0,9,9	-	-	-		
5	MGD	C	5102	7	43,52,52	1.48	8 (18%)	43,81,81	1.67	4 (9%)
9	EDO	G	7008	-	3,3,3	0.69	0	2,2,2	0.18	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
5	MGD	E	5201	7	-	3/18/66/66	0/6/6/6
9	EDO	A	7005	-	-	0/1/1/1	-
8	F3S	E	5205	1	-	-	0/3/3/3
9	EDO	C	7002	-	-	0/1/1/1	-
5	MGD	G	5301	7	-	4/18/66/66	0/6/6/6
9	EDO	G	7004	-	-	0/1/1/1	-
5	MGD	C	5101	7	-	5/18/66/66	0/6/6/6
5	MGD	A	5001	7	-	5/18/66/66	0/6/6/6
8	F3S	G	5305	1	-	-	0/3/3/3
5	MGD	G	5302	7	-	5/18/66/66	0/6/6/6
9	EDO	E	7003	-	-	0/1/1/1	-
10	FES	H	5306	2	-	-	0/1/1/1
8	F3S	A	5005	1	-	-	0/3/3/3
10	FES	D	5106	2	-	-	0/1/1/1
9	EDO	C	7006	-	-	0/1/1/1	-
9	EDO	A	7001	-	-	0/1/1/1	-
9	EDO	E	7007	-	-	0/1/1/1	-
5	MGD	E	5202	7	-	4/18/66/66	0/6/6/6
10	FES	B	5006	2	-	-	0/1/1/1
10	FES	F	5206	2	-	-	0/1/1/1
5	MGD	A	5002	7	-	5/18/66/66	0/6/6/6
9	EDO	E	7009	-	-	0/1/1/1	-
8	F3S	C	5105	1	-	-	0/3/3/3
5	MGD	C	5102	7	-	4/18/66/66	0/6/6/6
9	EDO	G	7008	-	-	0/1/1/1	-

The worst 5 of 73 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	5301	MGD	C23-C14	6.00	1.58	1.53
5	E	5201	MGD	C23-C14	5.00	1.57	1.53
5	C	5101	MGD	C23-C14	4.04	1.56	1.53
5	A	5002	MGD	C14-N15	3.94	1.50	1.46
5	A	5002	MGD	C23-C14	3.91	1.56	1.53

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	G	5301	MGD	C23-C14-N15	6.75	114.50	107.87
5	G	5302	MGD	C19-N20-C21	6.42	124.69	113.36
5	A	5002	MGD	C23-C14-N15	5.84	113.61	107.87
5	E	5202	MGD	C4'-O4'-C1'	-5.80	104.61	109.92
5	E	5202	MGD	C19-N20-C21	5.60	123.24	113.36

There are no chirality outliers.

5 of 35 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	5001	MGD	C5'-O5'-PB-O1B
5	A	5002	MGD	PA-O3B-PB-O5'
5	A	5002	MGD	C5'-O5'-PB-O1B
5	C	5101	MGD	PA-O3B-PB-O5'
5	C	5101	MGD	C5'-O5'-PB-O1B

There are no ring outliers.

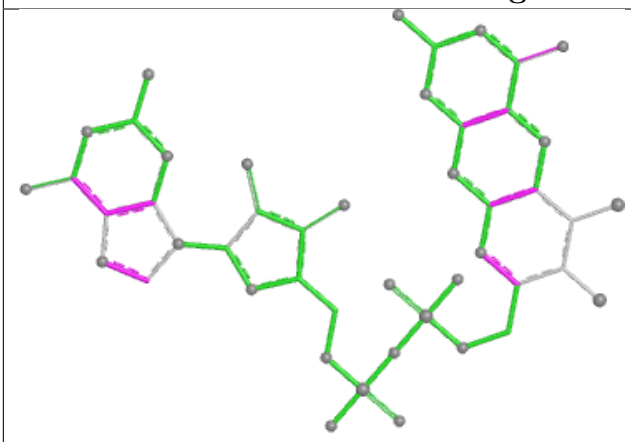
5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	5201	MGD	2	0
5	G	5301	MGD	1	0
5	C	5101	MGD	2	0
5	A	5001	MGD	1	0
5	A	5002	MGD	1	0

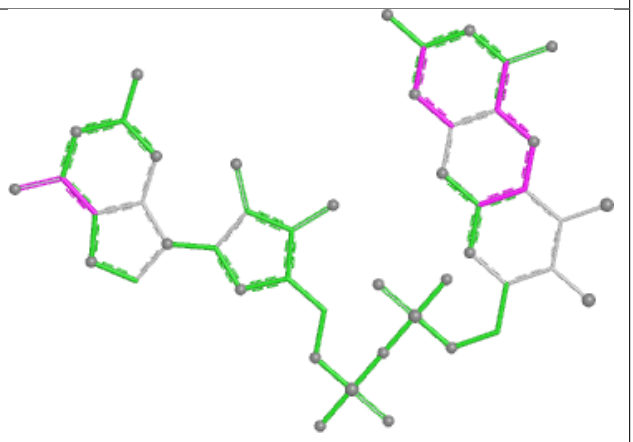
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

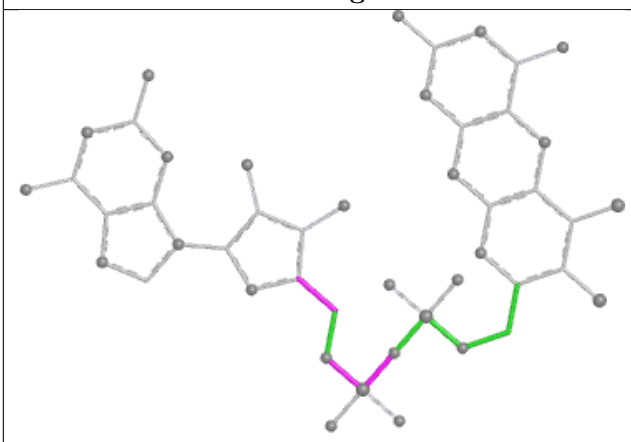
Ligand MGD E 5201



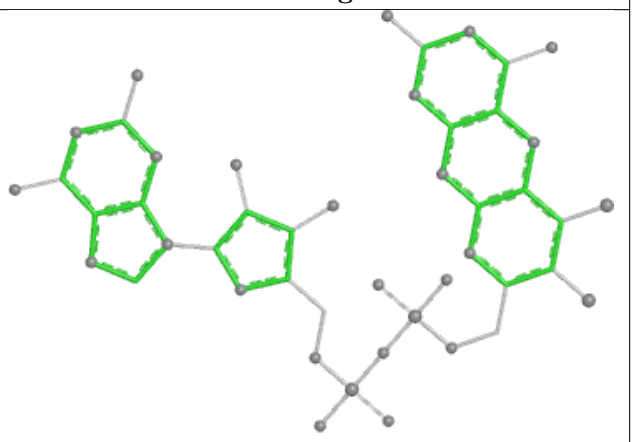
Bond lengths



Bond angles

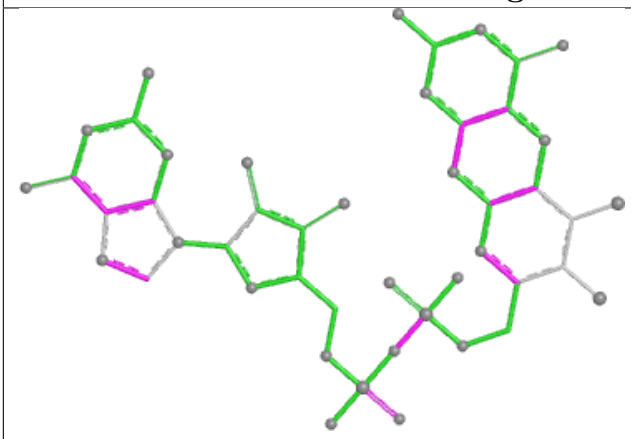


Torsions

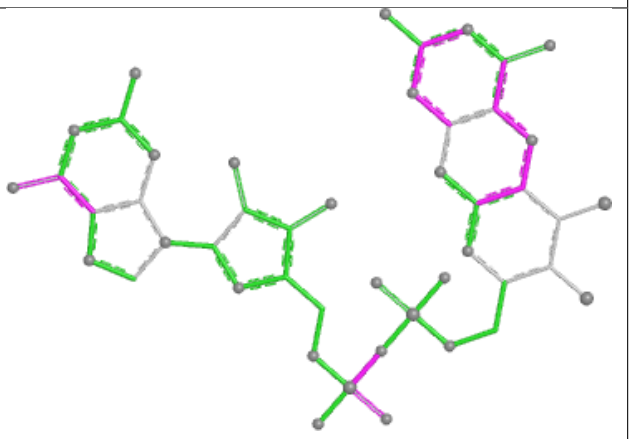


Rings

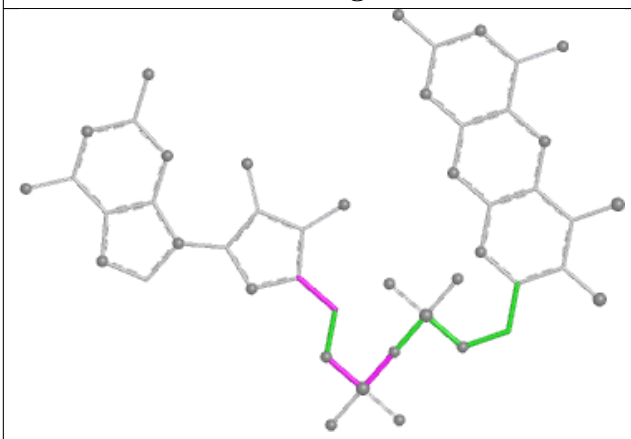
Ligand MGD G 5301



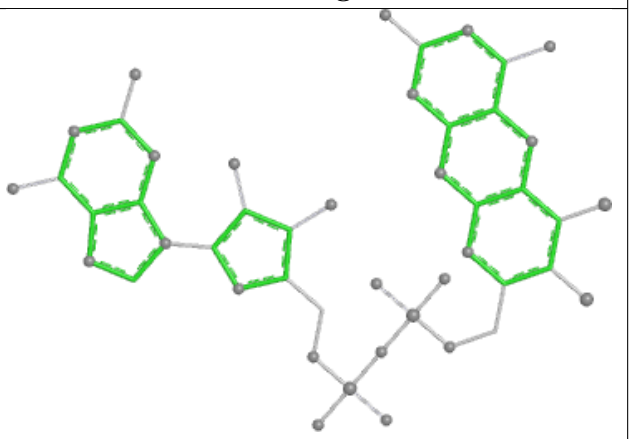
Bond lengths



Bond angles

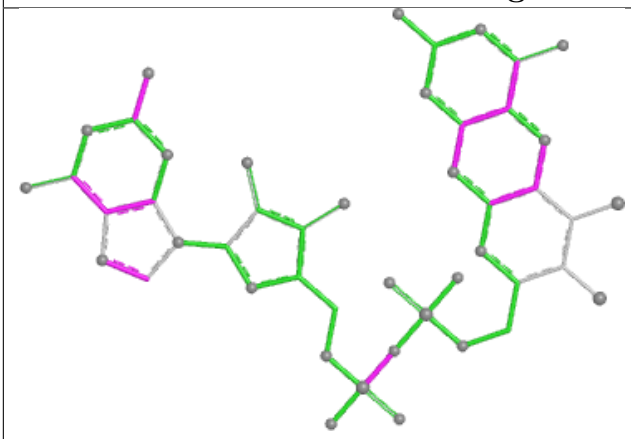


Torsions

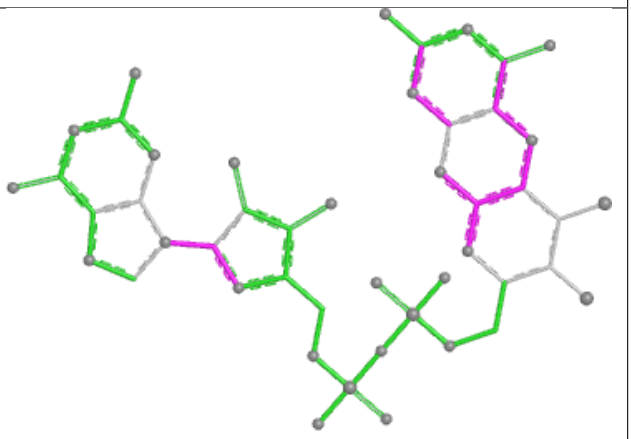


Rings

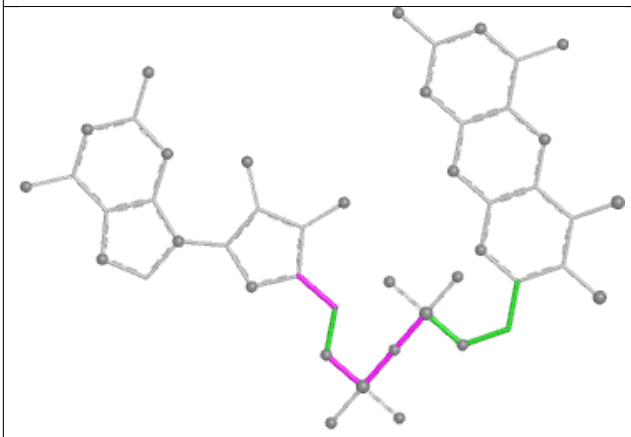
Ligand MGD C 5101



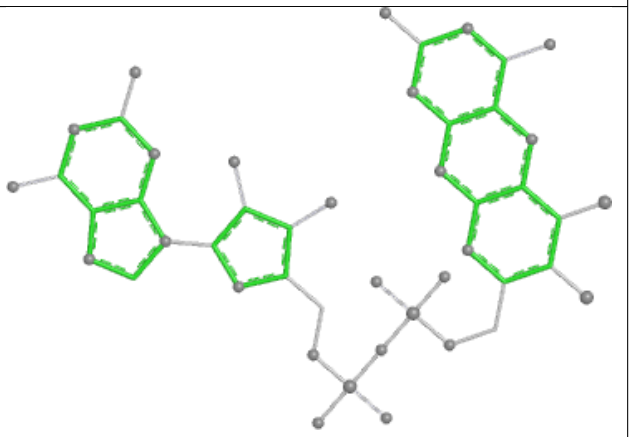
Bond lengths



Bond angles

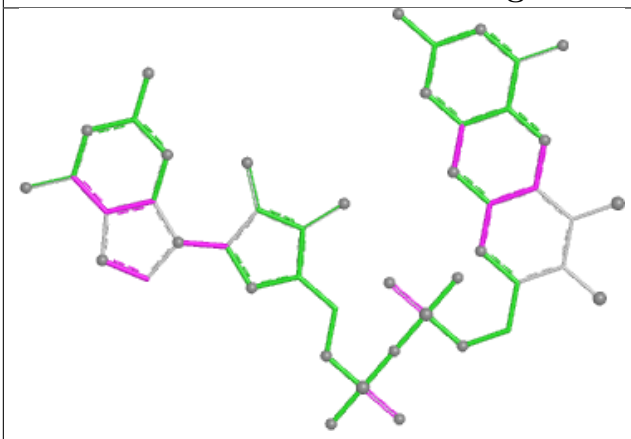


Torsions

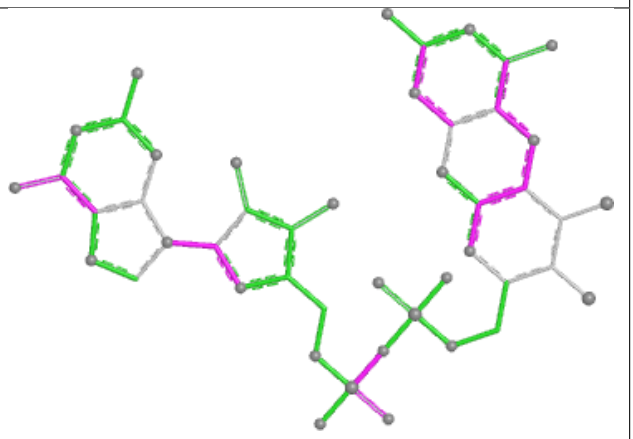


Rings

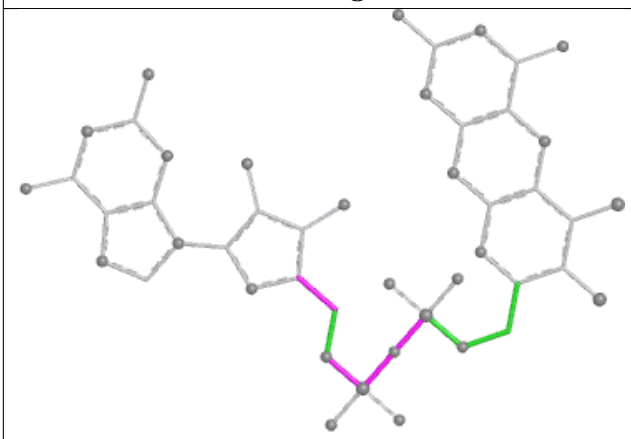
Ligand MGD A 5001



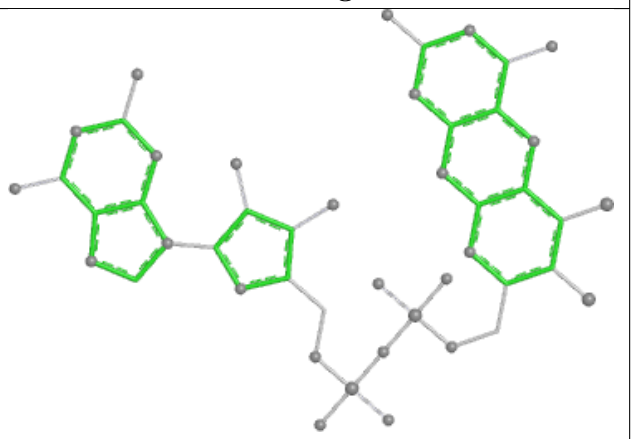
Bond lengths



Bond angles

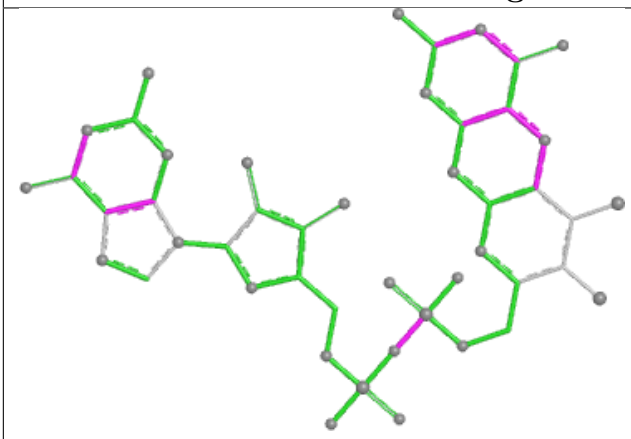


Torsions

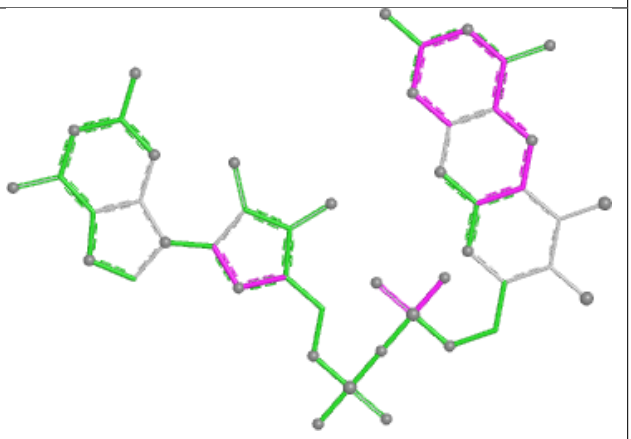


Rings

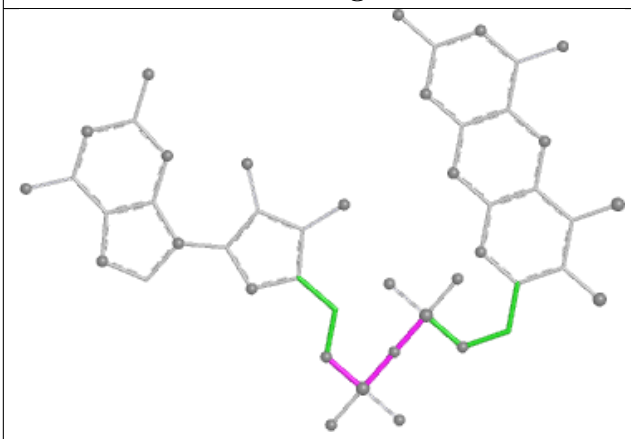
Ligand MGD G 5302



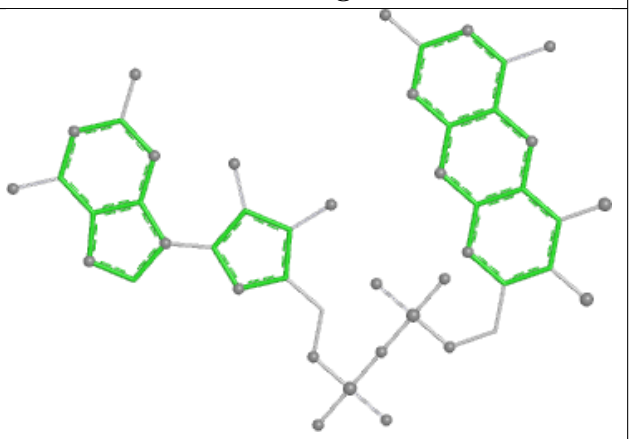
Bond lengths



Bond angles

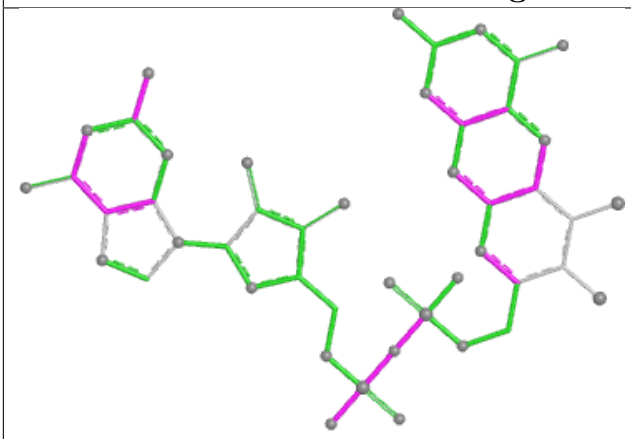


Torsions

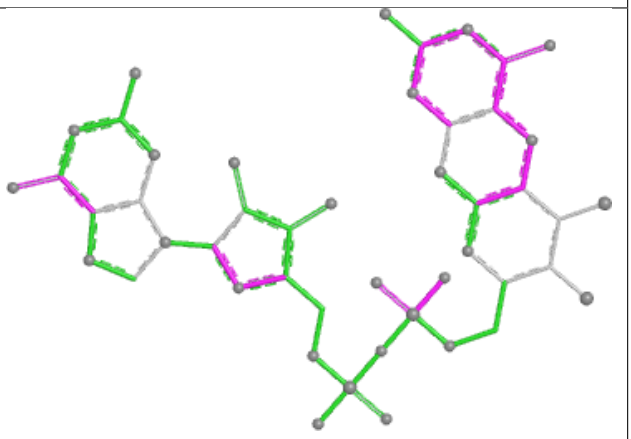


Rings

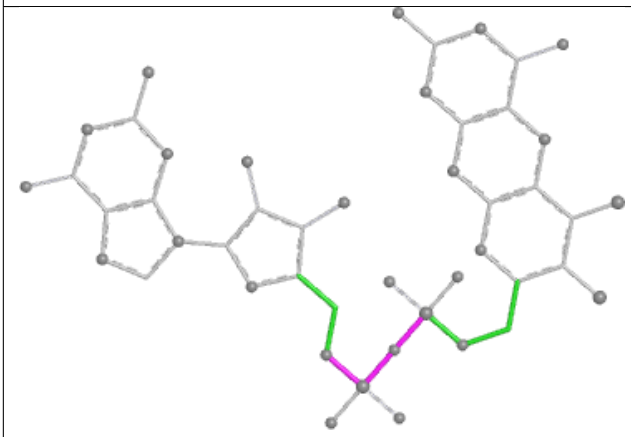
Ligand MGD E 5202



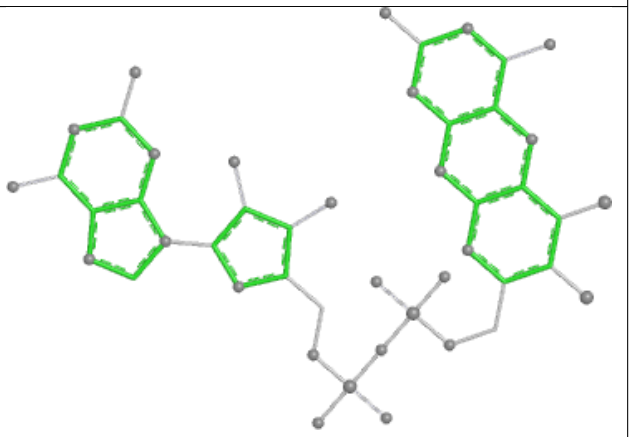
Bond lengths



Bond angles

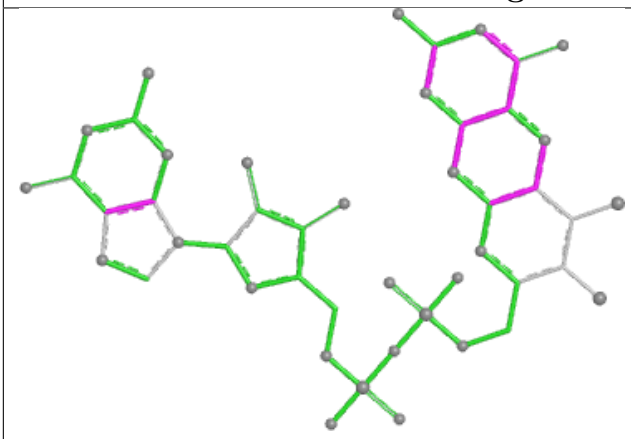


Torsions

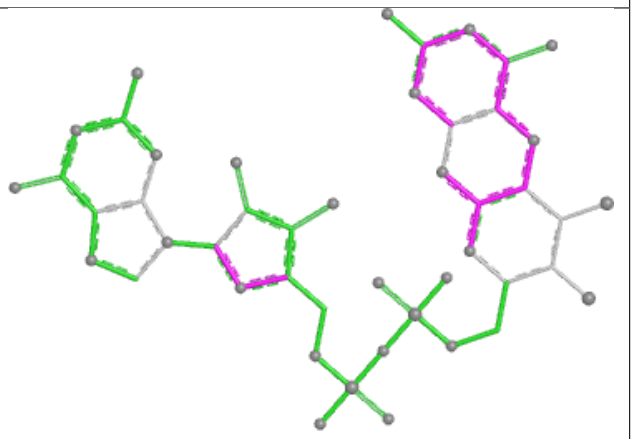


Rings

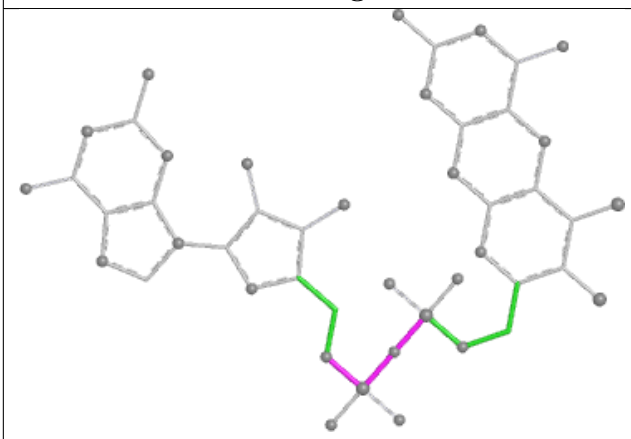
Ligand MGD A 5002



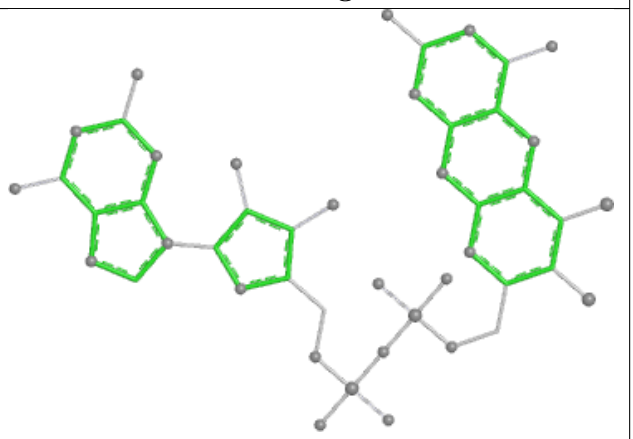
Bond lengths



Bond angles

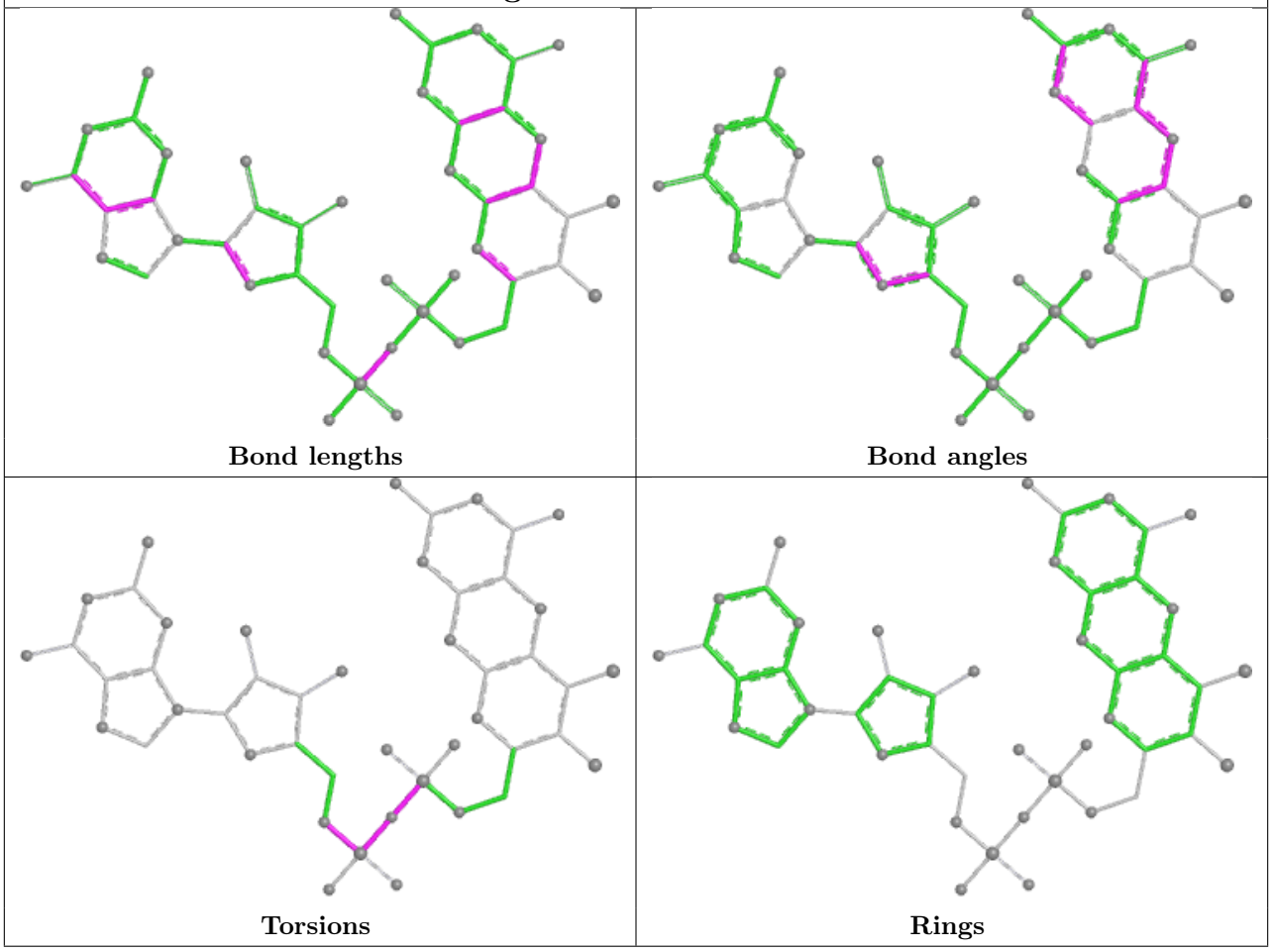


Torsions



Rings

Ligand MGD C 5102



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	822/825 (99%)	-0.58	0 100 100	9, 15, 24, 36	26 (3%)
1	C	822/825 (99%)	-0.73	1 (0%) 92 93	8, 13, 22, 30	23 (2%)
1	E	822/825 (99%)	-0.75	0 100 100	8, 13, 21, 31	28 (3%)
1	G	822/825 (99%)	-0.65	1 (0%) 92 93	9, 15, 23, 32	29 (3%)
2	B	133/133 (100%)	-0.66	0 100 100	10, 14, 20, 27	0
2	D	133/133 (100%)	-0.57	0 100 100	10, 14, 22, 28	1 (0%)
2	F	133/133 (100%)	-0.58	0 100 100	10, 14, 21, 27	1 (0%)
2	H	133/133 (100%)	-0.65	1 (0%) 82 85	10, 14, 20, 27	1 (0%)
All	All	3820/3832 (99%)	-0.67	3 (0%) 92 93	8, 14, 23, 36	109 (2%)

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	4	ASN	2.8
1	G	4	ASN	2.5
2	H	45	GLY	2.2

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

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, 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
3	HG	A	6009	1/1	0.95	0.09	29,29,29,29	1
9	EDO	A	7001	4/4	0.95	0.07	13,15,15,16	0
9	EDO	E	7009	4/4	0.95	0.07	14,16,17,18	0
3	HG	E	6011	1/1	0.96	0.08	29,29,29,29	1
9	EDO	E	7003	4/4	0.97	0.06	11,12,12,14	0
9	EDO	G	7004	4/4	0.97	0.06	12,13,13,15	0
6	O	E	5203	1/1	0.98	0.04	9,9,9,9	0
3	HG	C	6006	1/1	0.98	0.07	15,15,15,15	1
9	EDO	A	7005	4/4	0.98	0.04	12,14,14,15	0
9	EDO	C	7002	4/4	0.98	0.05	9,10,11,12	0
3	HG	C	6010	1/1	0.98	0.07	32,32,32,32	1
9	EDO	E	7007	4/4	0.98	0.04	11,12,12,14	0
3	HG	A	6005	1/1	0.98	0.07	18,18,18,18	1
6	O	A	5003	1/1	0.98	0.06	15,15,15,15	0
9	EDO	G	7008	4/4	0.98	0.04	13,14,14,15	0
5	MGD	A	5002	47/47	0.99	0.03	9,10,12,12	0
5	MGD	C	5101	47/47	0.99	0.04	8,10,12,12	0
5	MGD	C	5102	47/47	0.99	0.04	7,9,10,11	0
5	MGD	E	5201	47/47	0.99	0.03	8,9,10,11	0
5	MGD	E	5202	47/47	0.99	0.03	6,8,9,10	0
5	MGD	G	5301	47/47	0.99	0.03	7,10,12,12	0
5	MGD	G	5302	47/47	0.99	0.03	7,10,11,12	0
3	HG	E	6007	1/1	0.99	0.08	15,15,15,15	1
6	O	C	5103	1/1	0.99	0.04	10,10,10,10	0
3	HG	B	6001	1/1	0.99	0.06	14,14,14,14	1
3	HG	F	6003	1/1	0.99	0.06	13,13,13,13	1
3	HG	G	6008	1/1	0.99	0.07	17,17,17,17	1
3	HG	G	6012	1/1	0.99	0.08	30,30,30,30	1
9	EDO	C	7006	4/4	0.99	0.04	9,10,11,11	0
4	CA	C	5107	1/1	0.99	0.05	19,19,19,19	0
4	CA	C	5108	1/1	0.99	0.04	14,14,14,14	0
4	CA	G	5307	1/1	0.99	0.04	18,18,18,18	0
4	CA	G	5308	1/1	0.99	0.03	13,13,13,13	0
5	MGD	A	5001	47/47	0.99	0.04	10,11,13,13	0
7	4MO	E	5204	1/1	1.00	0.02	10,10,10,10	0
7	4MO	G	5304	1/1	1.00	0.02	12,12,12,12	0
8	F3S	A	5005	7/7	1.00	0.02	9,10,11,11	0

Continued on next page...

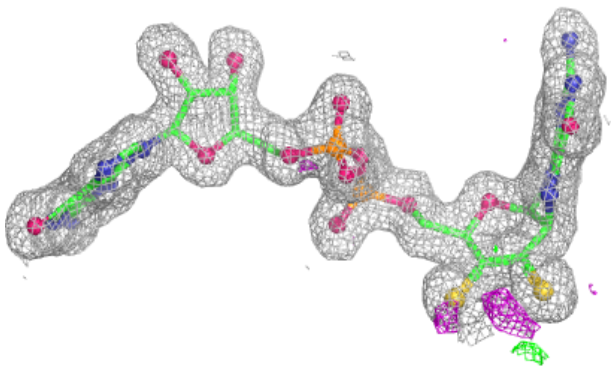
Continued from previous page...

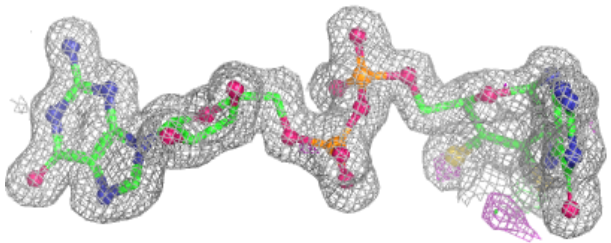
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
8	F3S	C	5105	7/7	1.00	0.01	6,8,9,10	0
8	F3S	E	5205	7/7	1.00	0.01	8,8,8,9	0
8	F3S	G	5305	7/7	1.00	0.02	9,10,10,10	0
4	CA	A	5008	1/1	1.00	0.04	12,12,12,12	0
3	HG	D	6002	1/1	1.00	0.07	14,14,14,14	1
3	HG	H	6004	1/1	1.00	0.06	13,13,13,13	1
4	CA	E	5207	1/1	1.00	0.04	18,18,18,18	0
4	CA	E	5208	1/1	1.00	0.03	12,12,12,12	0
4	CA	A	5007	1/1	1.00	0.04	18,18,18,18	0
6	O	G	5303	1/1	1.00	0.03	12,12,12,12	0
7	4MO	A	5004	1/1	1.00	0.01	12,12,12,12	0
7	4MO	C	5104	1/1	1.00	0.01	10,10,10,10	0
10	FES	B	5006	4/4	1.00	0.02	11,11,12,12	0
10	FES	D	5106	4/4	1.00	0.01	11,11,12,12	0
10	FES	F	5206	4/4	1.00	0.01	11,11,11,11	0
10	FES	H	5306	4/4	1.00	0.02	11,12,12,13	0

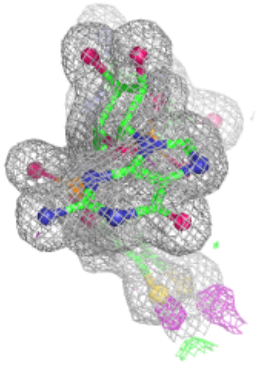
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around MGD A 5002:

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

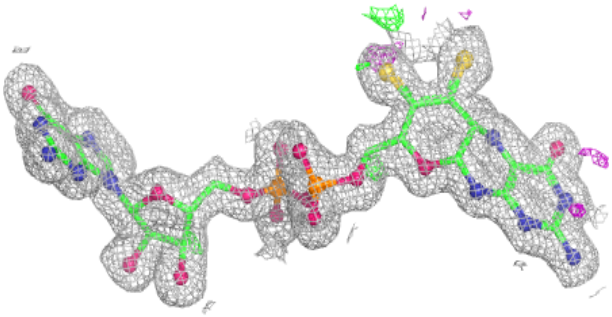


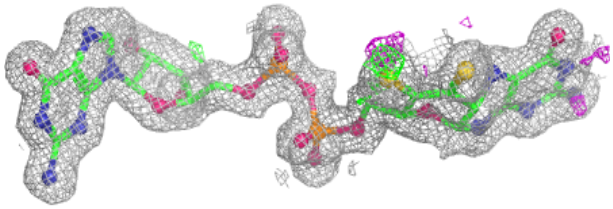


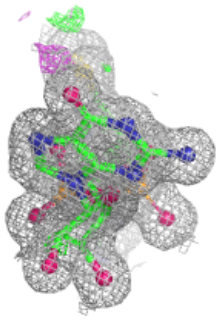


Electron density around MGD C 5101:

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

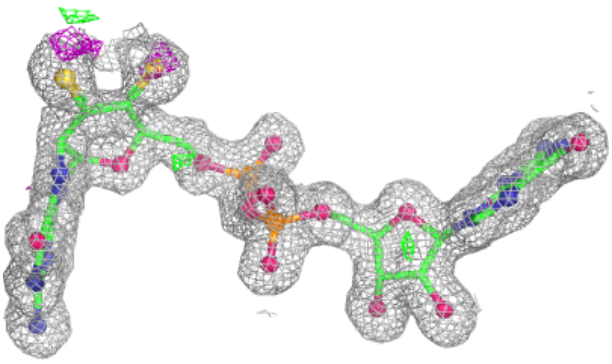


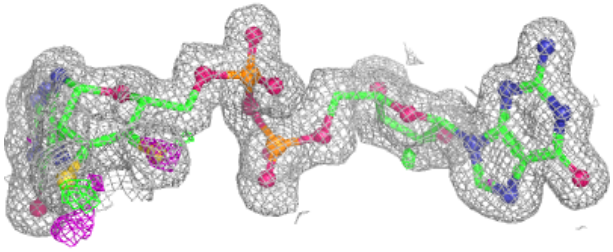


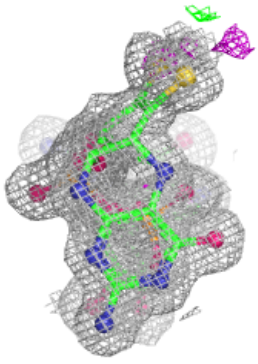


Electron density around MGD C 5102:

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

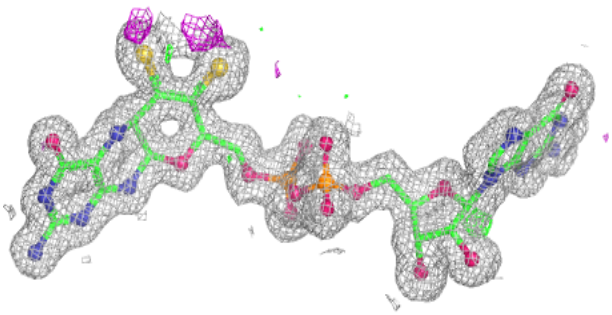


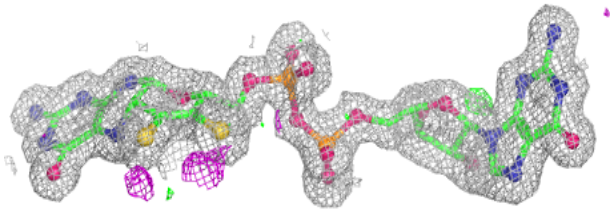


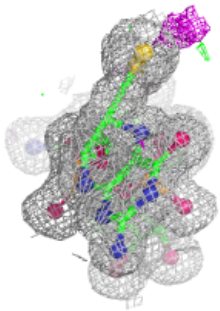


Electron density around MGD E 5201:

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

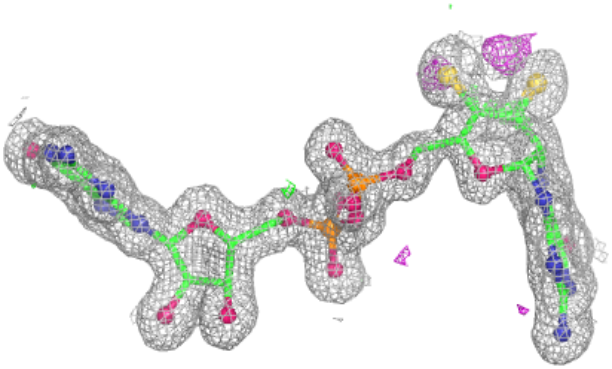


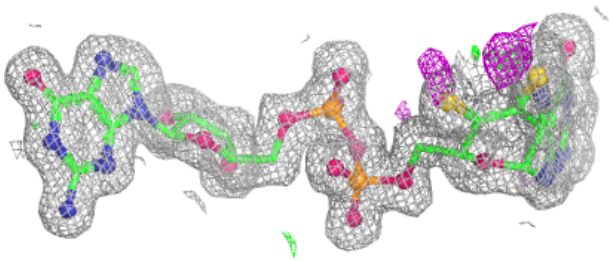


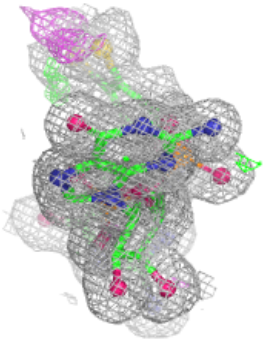


Electron density around MGD E 5202:

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

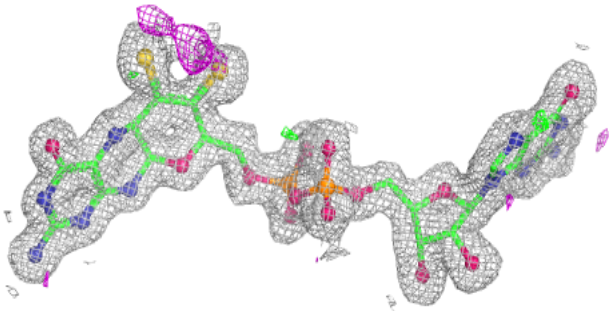


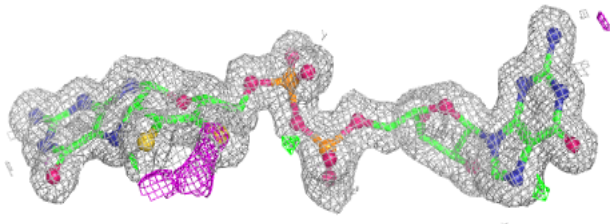


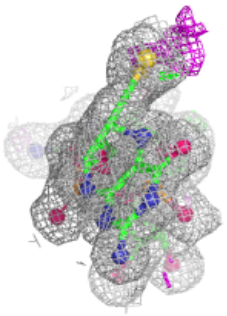


Electron density around MGD G 5301:

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

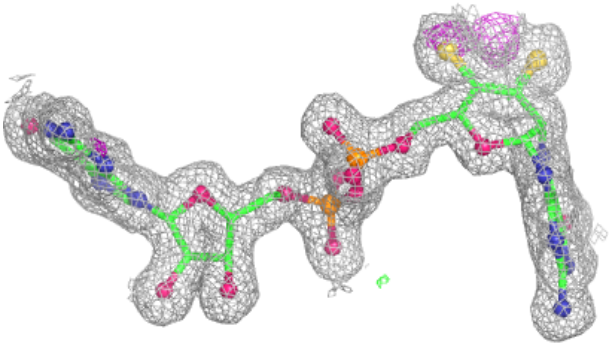


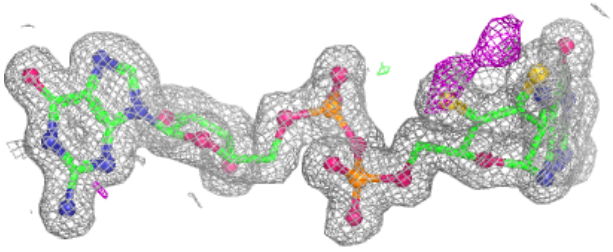


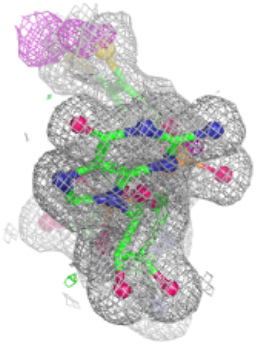


Electron density around MGD G 5302:

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

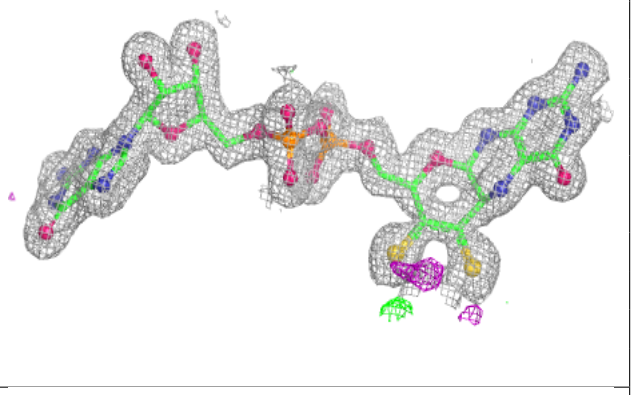


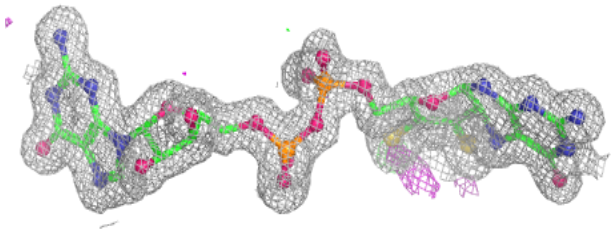


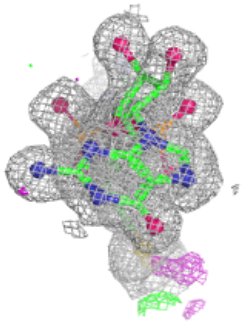


Electron density around MGD A 5001:

$2mF_o$-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.