wwPDB X-ray Structure Validation Summary Report (i)

Dec 6, 2021 – 01:07 pm GMT

PDB ID	:	7PCX
Title	:	X-RAY STRUCTURE OF THE HALOALKANE DEHALOGENASE HALOTAG7-Q165W LABELED WITH A CHLOROALKANE-TETRAM ETHYLRHODAMINE FLUOROPHORE SUBSTRATE
Authors	:	Tarnawski, M.; Frei, M.; Hiblot, J.; Johnsson, K.
Deposited on	:	2021-08-04
Resolution	:	1.40 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:

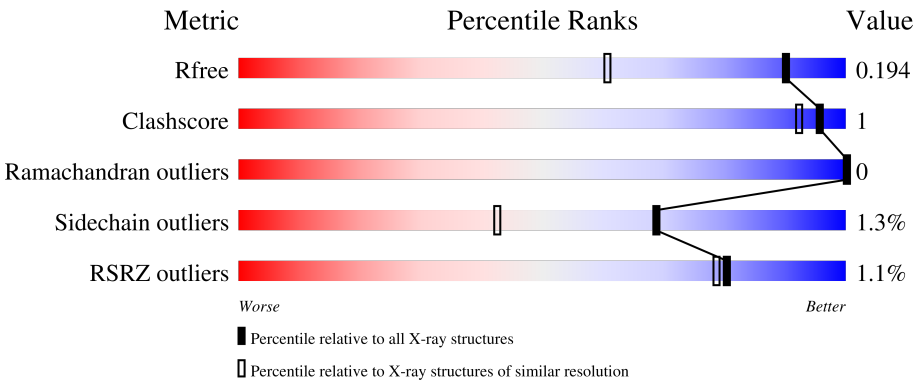
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

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

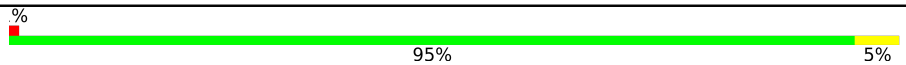
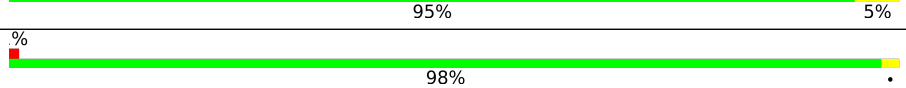
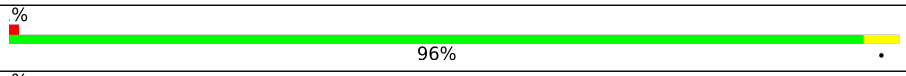
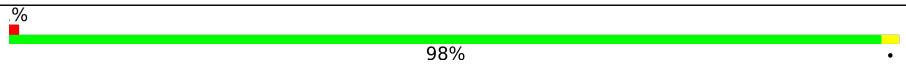
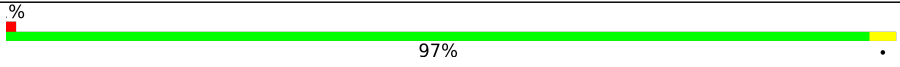
The reported resolution of this entry is 1.40 Å.

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	1714 (1.40-1.40)
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)
RSRZ outliers	127900	1674 (1.40-1.40)

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	293	
1	B	293	
1	C	293	
1	D	293	
1	E	293	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	293	% 96% .

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 15752 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 Haloalkane dehalogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	293	2366	1543	395	419	9	0	2	0
1	B	293	2363	1539	398	417	9	0	1	0
1	C	293	2366	1543	395	419	9	0	2	0
1	D	293	2388	1559	395	425	9	0	6	0
1	E	293	2377	1551	396	421	9	0	3	0
1	F	293	2377	1552	395	421	9	0	4	0

There are 150 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	GLY	-	expression tag	UNP P0A3G3
A	47	VAL	LEU	engineered mutation	UNP P0A3G3
A	58	THR	SER	engineered mutation	UNP P0A3G3
A	78	GLY	ASP	engineered mutation	UNP P0A3G3
A	87	PHE	TYR	engineered mutation	UNP P0A3G3
A	88	MET	LEU	engineered mutation	UNP P0A3G3
A	128	PHE	CYS	engineered mutation	UNP P0A3G3
A	155	THR	ALA	engineered mutation	UNP P0A3G3
A	160	LYS	GLU	engineered mutation	UNP P0A3G3
A	165	TRP	GLN	engineered mutation	UNP P0A3G3
A	167	VAL	ALA	engineered mutation	UNP P0A3G3
A	172	THR	ALA	engineered mutation	UNP P0A3G3
A	175	MET	LYS	engineered mutation	UNP P0A3G3
A	176	GLY	CYS	engineered mutation	UNP P0A3G3
A	195	ASN	LYS	engineered mutation	UNP P0A3G3
A	224	GLU	ALA	engineered mutation	UNP P0A3G3
A	227	ASP	ASN	engineered mutation	UNP P0A3G3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	257	LYS	GLU	engineered mutation	UNP P0A3G3
A	264	ALA	THR	engineered mutation	UNP P0A3G3
A	272	ASN	HIS	engineered mutation	UNP P0A3G3
A	273	LEU	TYR	engineered mutation	UNP P0A3G3
A	291	SER	PRO	engineered mutation	UNP P0A3G3
A	292	THR	ALA	engineered mutation	UNP P0A3G3
A	294	GLU	-	expression tag	UNP P0A3G3
A	295	ILE	-	expression tag	UNP P0A3G3
B	3	GLY	-	expression tag	UNP P0A3G3
B	47	VAL	LEU	engineered mutation	UNP P0A3G3
B	58	THR	SER	engineered mutation	UNP P0A3G3
B	78	GLY	ASP	engineered mutation	UNP P0A3G3
B	87	PHE	TYR	engineered mutation	UNP P0A3G3
B	88	MET	LEU	engineered mutation	UNP P0A3G3
B	128	PHE	CYS	engineered mutation	UNP P0A3G3
B	155	THR	ALA	engineered mutation	UNP P0A3G3
B	160	LYS	GLU	engineered mutation	UNP P0A3G3
B	165	TRP	GLN	engineered mutation	UNP P0A3G3
B	167	VAL	ALA	engineered mutation	UNP P0A3G3
B	172	THR	ALA	engineered mutation	UNP P0A3G3
B	175	MET	LYS	engineered mutation	UNP P0A3G3
B	176	GLY	CYS	engineered mutation	UNP P0A3G3
B	195	ASN	LYS	engineered mutation	UNP P0A3G3
B	224	GLU	ALA	engineered mutation	UNP P0A3G3
B	227	ASP	ASN	engineered mutation	UNP P0A3G3
B	257	LYS	GLU	engineered mutation	UNP P0A3G3
B	264	ALA	THR	engineered mutation	UNP P0A3G3
B	272	ASN	HIS	engineered mutation	UNP P0A3G3
B	273	LEU	TYR	engineered mutation	UNP P0A3G3
B	291	SER	PRO	engineered mutation	UNP P0A3G3
B	292	THR	ALA	engineered mutation	UNP P0A3G3
B	294	GLU	-	expression tag	UNP P0A3G3
B	295	ILE	-	expression tag	UNP P0A3G3
C	3	GLY	-	expression tag	UNP P0A3G3
C	47	VAL	LEU	engineered mutation	UNP P0A3G3
C	58	THR	SER	engineered mutation	UNP P0A3G3
C	78	GLY	ASP	engineered mutation	UNP P0A3G3
C	87	PHE	TYR	engineered mutation	UNP P0A3G3
C	88	MET	LEU	engineered mutation	UNP P0A3G3
C	128	PHE	CYS	engineered mutation	UNP P0A3G3
C	155	THR	ALA	engineered mutation	UNP P0A3G3
C	160	LYS	GLU	engineered mutation	UNP P0A3G3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	165	TRP	GLN	engineered mutation	UNP P0A3G3
C	167	VAL	ALA	engineered mutation	UNP P0A3G3
C	172	THR	ALA	engineered mutation	UNP P0A3G3
C	175	MET	LYS	engineered mutation	UNP P0A3G3
C	176	GLY	CYS	engineered mutation	UNP P0A3G3
C	195	ASN	LYS	engineered mutation	UNP P0A3G3
C	224	GLU	ALA	engineered mutation	UNP P0A3G3
C	227	ASP	ASN	engineered mutation	UNP P0A3G3
C	257	LYS	GLU	engineered mutation	UNP P0A3G3
C	264	ALA	THR	engineered mutation	UNP P0A3G3
C	272	ASN	HIS	engineered mutation	UNP P0A3G3
C	273	LEU	TYR	engineered mutation	UNP P0A3G3
C	291	SER	PRO	engineered mutation	UNP P0A3G3
C	292	THR	ALA	engineered mutation	UNP P0A3G3
C	294	GLU	-	expression tag	UNP P0A3G3
C	295	ILE	-	expression tag	UNP P0A3G3
D	3	GLY	-	expression tag	UNP P0A3G3
D	47	VAL	LEU	engineered mutation	UNP P0A3G3
D	58	THR	SER	engineered mutation	UNP P0A3G3
D	78	GLY	ASP	engineered mutation	UNP P0A3G3
D	87	PHE	TYR	engineered mutation	UNP P0A3G3
D	88	MET	LEU	engineered mutation	UNP P0A3G3
D	128	PHE	CYS	engineered mutation	UNP P0A3G3
D	155	THR	ALA	engineered mutation	UNP P0A3G3
D	160	LYS	GLU	engineered mutation	UNP P0A3G3
D	165	TRP	GLN	engineered mutation	UNP P0A3G3
D	167	VAL	ALA	engineered mutation	UNP P0A3G3
D	172	THR	ALA	engineered mutation	UNP P0A3G3
D	175	MET	LYS	engineered mutation	UNP P0A3G3
D	176	GLY	CYS	engineered mutation	UNP P0A3G3
D	195	ASN	LYS	engineered mutation	UNP P0A3G3
D	224	GLU	ALA	engineered mutation	UNP P0A3G3
D	227	ASP	ASN	engineered mutation	UNP P0A3G3
D	257	LYS	GLU	engineered mutation	UNP P0A3G3
D	264	ALA	THR	engineered mutation	UNP P0A3G3
D	272	ASN	HIS	engineered mutation	UNP P0A3G3
D	273	LEU	TYR	engineered mutation	UNP P0A3G3
D	291	SER	PRO	engineered mutation	UNP P0A3G3
D	292	THR	ALA	engineered mutation	UNP P0A3G3
D	294	GLU	-	expression tag	UNP P0A3G3
D	295	ILE	-	expression tag	UNP P0A3G3
E	3	GLY	-	expression tag	UNP P0A3G3

Continued on next page...

Continued from previous page...

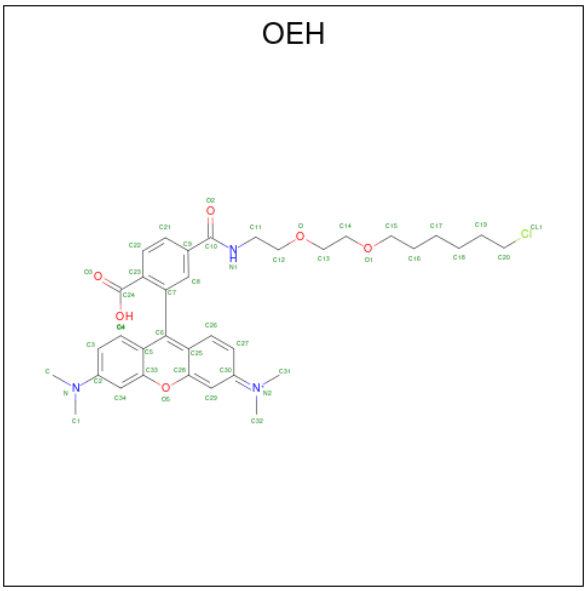
Chain	Residue	Modelled	Actual	Comment	Reference
E	47	VAL	LEU	engineered mutation	UNP P0A3G3
E	58	THR	SER	engineered mutation	UNP P0A3G3
E	78	GLY	ASP	engineered mutation	UNP P0A3G3
E	87	PHE	TYR	engineered mutation	UNP P0A3G3
E	88	MET	LEU	engineered mutation	UNP P0A3G3
E	128	PHE	CYS	engineered mutation	UNP P0A3G3
E	155	THR	ALA	engineered mutation	UNP P0A3G3
E	160	LYS	GLU	engineered mutation	UNP P0A3G3
E	165	TRP	GLN	engineered mutation	UNP P0A3G3
E	167	VAL	ALA	engineered mutation	UNP P0A3G3
E	172	THR	ALA	engineered mutation	UNP P0A3G3
E	175	MET	LYS	engineered mutation	UNP P0A3G3
E	176	GLY	CYS	engineered mutation	UNP P0A3G3
E	195	ASN	LYS	engineered mutation	UNP P0A3G3
E	224	GLU	ALA	engineered mutation	UNP P0A3G3
E	227	ASP	ASN	engineered mutation	UNP P0A3G3
E	257	LYS	GLU	engineered mutation	UNP P0A3G3
E	264	ALA	THR	engineered mutation	UNP P0A3G3
E	272	ASN	HIS	engineered mutation	UNP P0A3G3
E	273	LEU	TYR	engineered mutation	UNP P0A3G3
E	291	SER	PRO	engineered mutation	UNP P0A3G3
E	292	THR	ALA	engineered mutation	UNP P0A3G3
E	294	GLU	-	expression tag	UNP P0A3G3
E	295	ILE	-	expression tag	UNP P0A3G3
F	3	GLY	-	expression tag	UNP P0A3G3
F	47	VAL	LEU	engineered mutation	UNP P0A3G3
F	58	THR	SER	engineered mutation	UNP P0A3G3
F	78	GLY	ASP	engineered mutation	UNP P0A3G3
F	87	PHE	TYR	engineered mutation	UNP P0A3G3
F	88	MET	LEU	engineered mutation	UNP P0A3G3
F	128	PHE	CYS	engineered mutation	UNP P0A3G3
F	155	THR	ALA	engineered mutation	UNP P0A3G3
F	160	LYS	GLU	engineered mutation	UNP P0A3G3
F	165	TRP	GLN	engineered mutation	UNP P0A3G3
F	167	VAL	ALA	engineered mutation	UNP P0A3G3
F	172	THR	ALA	engineered mutation	UNP P0A3G3
F	175	MET	LYS	engineered mutation	UNP P0A3G3
F	176	GLY	CYS	engineered mutation	UNP P0A3G3
F	195	ASN	LYS	engineered mutation	UNP P0A3G3
F	224	GLU	ALA	engineered mutation	UNP P0A3G3
F	227	ASP	ASN	engineered mutation	UNP P0A3G3
F	257	LYS	GLU	engineered mutation	UNP P0A3G3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	264	ALA	THR	engineered mutation	UNP P0A3G3
F	272	ASN	HIS	engineered mutation	UNP P0A3G3
F	273	LEU	TYR	engineered mutation	UNP P0A3G3
F	291	SER	PRO	engineered mutation	UNP P0A3G3
F	292	THR	ALA	engineered mutation	UNP P0A3G3
F	294	GLU	-	expression tag	UNP P0A3G3
F	295	ILE	-	expression tag	UNP P0A3G3

- Molecule 2 is [9-[2-carboxy-5-[2-[2-(6-chloranylhexoxy)ethoxy]ethylcarbamoyl]phenyl]-6-(dimethylamino)xanthen-3-ylidene]-dimethyl-azanium (three-letter code: OEH) (formula: $C_{35}H_{43}ClN_3O_6$) (labeled as "Ligand of Interest" by depositor).

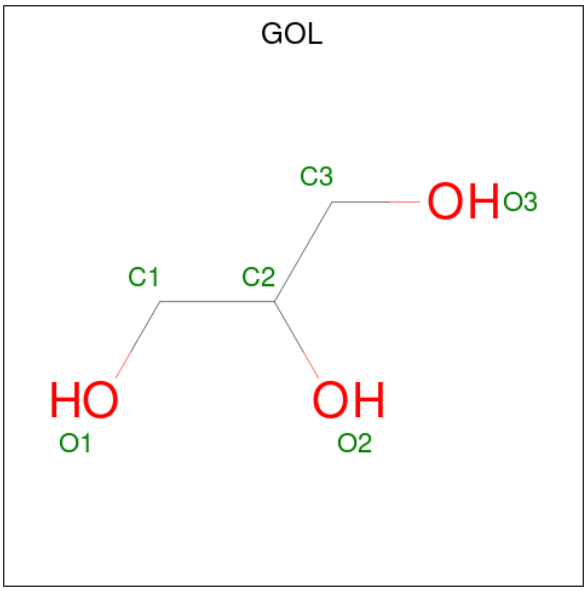


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	44	35	3	6	0	0
2	B	1	44	35	3	6	0	0
2	C	1	88	70	6	12	0	1
2	D	1	44	35	3	6	0	0
2	E	1	44	35	3	6	0	0
2	F	1	44	35	3	6	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0
3	B	1	Total Cl 1 1	0	0
3	C	1	Total Cl 1 1	0	0
3	D	1	Total Cl 1 1	0	0
3	E	1	Total Cl 1 1	0	0
3	F	1	Total Cl 1 1	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

Continued on next page...

Continued from previous page...

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

- Molecule 5 is water.

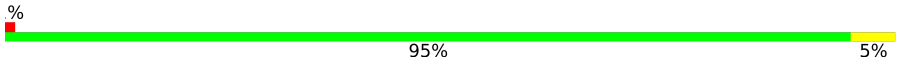
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	197	197	197	0	0
			Total	O		
5	B	200	200	200	0	0
			Total	O		
5	C	180	180	180	0	0
			Total	O		
5	D	202	202	202	0	0
			Total	O		
5	E	182	182	182	0	0
			Total	O		
5	F	198	198	198	0	0

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.

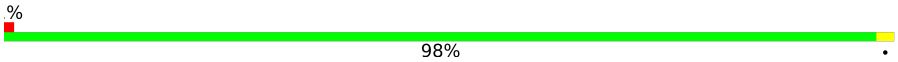
- Molecule 1: Haloalkane dehalogenase

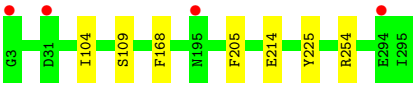
Chain A:



- Molecule 1: Haloalkane dehalogenase

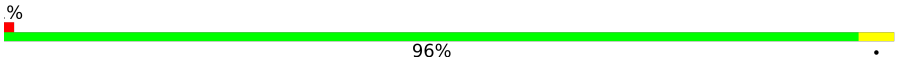
Chain B:

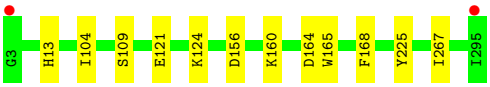




- Molecule 1: Haloalkane dehalogenase

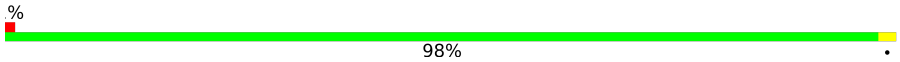
Chain C:





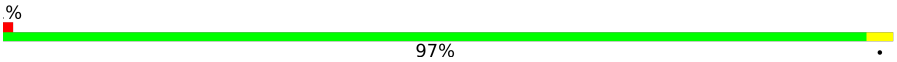
- Molecule 1: Haloalkane dehalogenase

Chain D:



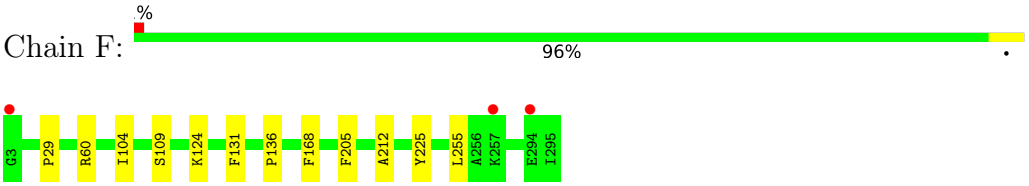
- Molecule 1: Haloalkane dehalogenase

Chain E:



- Molecule 1: Haloalkane dehalogenase

Chain F:


%
96%
G3
P29
R60
I104
S109
K124
F131
P136
F168
F205
A212
Y225
L255
A256
K257
E294
I295

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	74.96Å 149.15Å 79.92Å 90.00° 117.97° 90.00°	Depositor
Resolution (Å)	39.86 – 1.40 39.86 – 1.40	Depositor EDS
% Data completeness (in resolution range)	99.1 (39.86-1.40) 99.1 (39.86-1.40)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.96 (at 1.40Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.175 , 0.198 0.172 , 0.194	Depositor DCC
R_{free} test set	15027 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	11.7	Xtriage
Anisotropy	0.186	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.307 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	15752	wwPDB-VP
Average B, all atoms (Å^2)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 73.40 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.8505e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: CL, OEH, GOL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\|>5$	RMSZ	$\#\|Z\|>5$
1	A	0.34	0/2453	0.61	0/3355
1	B	0.35	0/2447	0.63	0/3345
1	C	0.35	0/2453	0.62	1/3356 (0.0%)
1	D	0.34	0/2487	0.61	0/3401
1	E	0.36	0/2469	0.62	0/3377
1	F	0.35	0/2470	0.61	0/3378
All	All	0.35	0/14779	0.62	1/20212 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	156	ASP	CB-CG-OD1	5.10	122.89	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	2366	0	2310	7	0
1	B	2363	0	2306	5	0
1	C	2366	0	2310	7	0
1	D	2388	0	2337	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	2377	0	2313	6	0
1	F	2377	0	2327	6	0
2	A	44	0	0	0	0
2	B	44	0	0	0	0
2	C	88	0	0	0	0
2	D	44	0	0	0	0
2	E	44	0	0	1	0
2	F	44	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	6	0	8	0	0
4	B	6	0	8	0	0
4	C	6	0	8	0	0
4	D	6	0	8	0	0
4	E	12	0	16	0	0
4	F	6	0	8	0	0
5	A	197	0	0	0	0
5	B	200	0	0	0	0
5	C	180	0	0	0	0
5	D	202	0	0	0	0
5	E	182	0	0	1	0
5	F	198	0	0	0	0
All	All	15752	0	13959	30	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 1.

The worst 5 of 30 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:B:254:ARG:HD3	1:C:13:HIS:HE1	1.50	0.75
1:A:82:ASP:OD2	1:A:86:ARG:NH1	2.27	0.68
1:D:265:VAL:HG21	1:D:285[B]:GLU:HG2	1.83	0.60
1:B:254:ARG:HD3	1:C:13:HIS:CE1	2.38	0.53
1:E:121[B]:GLU:HG2	1:E:122:ARG:HG3	1.93	0.49

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	293/293 (100%)	283 (97%)	10 (3%)	0	100	100
1	B	292/293 (100%)	282 (97%)	10 (3%)	0	100	100
1	C	293/293 (100%)	284 (97%)	9 (3%)	0	100	100
1	D	297/293 (101%)	288 (97%)	9 (3%)	0	100	100
1	E	294/293 (100%)	285 (97%)	9 (3%)	0	100	100
1	F	295/293 (101%)	286 (97%)	9 (3%)	0	100	100
All	All	1764/1758 (100%)	1708 (97%)	56 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	254/252 (101%)	250 (98%)	4 (2%)	62	33
1	B	253/252 (100%)	250 (99%)	3 (1%)	71	47
1	C	254/252 (101%)	251 (99%)	3 (1%)	71	47
1	D	258/252 (102%)	255 (99%)	3 (1%)	71	47
1	E	255/252 (101%)	252 (99%)	3 (1%)	71	47
1	F	256/252 (102%)	253 (99%)	3 (1%)	71	47
All	All	1530/1512 (101%)	1511 (99%)	19 (1%)	69	47

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	205	PHE
1	F	205	PHE
1	F	225	TYR
1	F	168	PHE
1	C	225	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	13	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 20 ligands modelled in this entry, 6 are monoatomic - leaving 14 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.

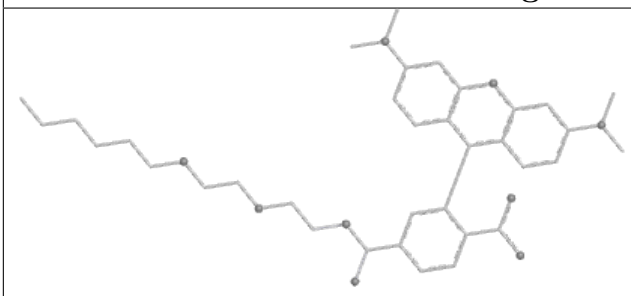
There are no ring outliers.

No monomer is involved in short contacts.

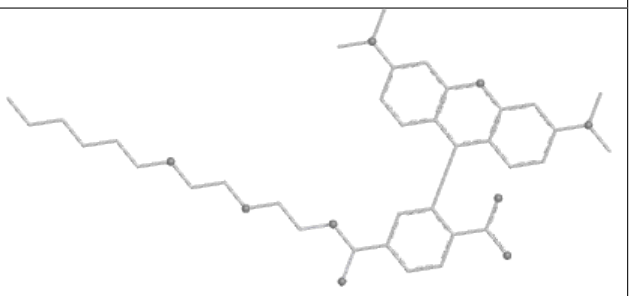
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

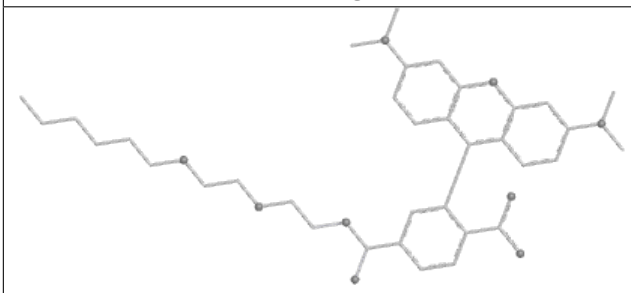
Ligand OEH A 301



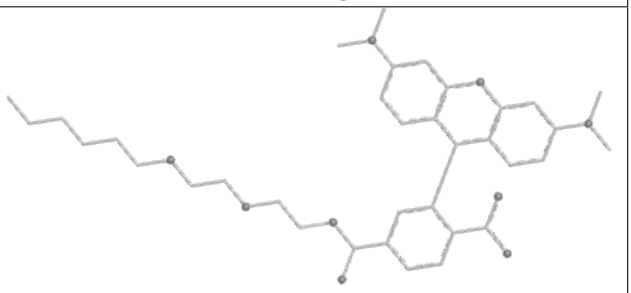
Bond lengths



Bond angles

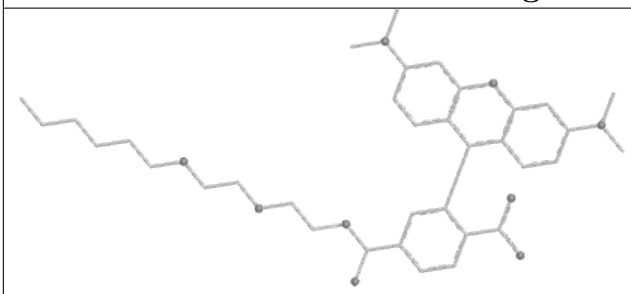


Torsions

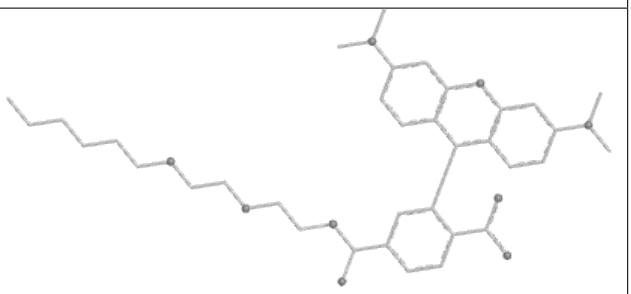


Rings

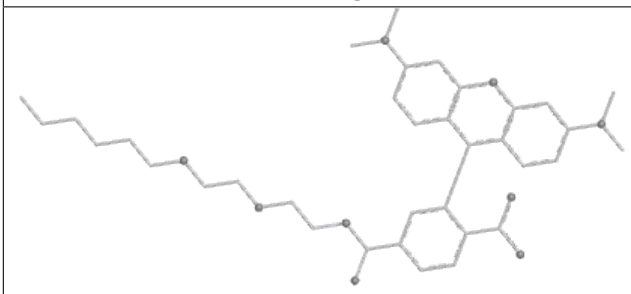
Ligand OEH B 301



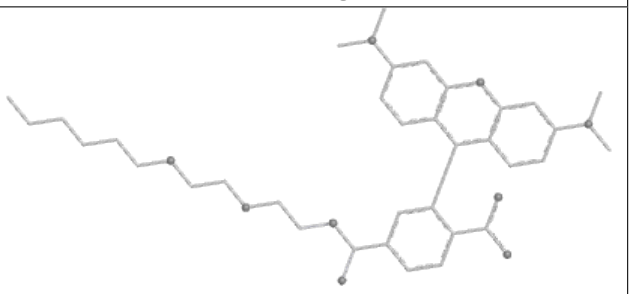
Bond lengths



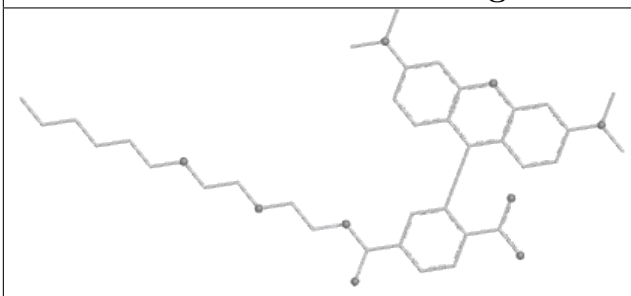
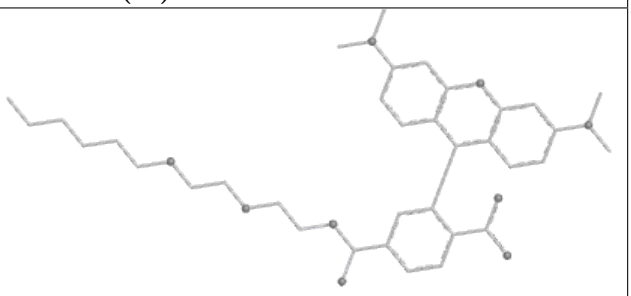
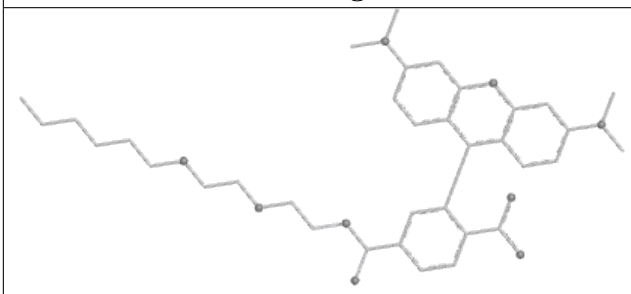
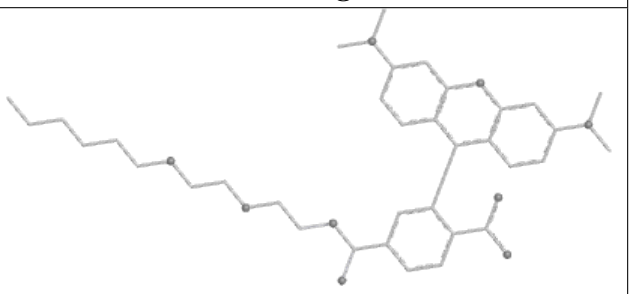
Bond angles

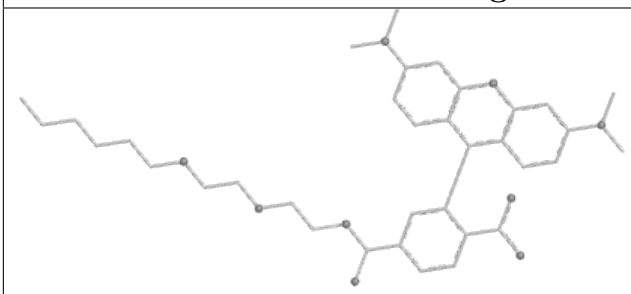
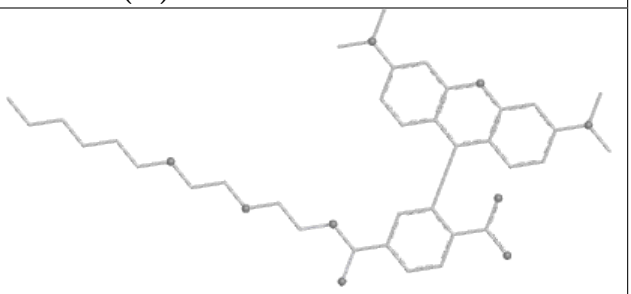
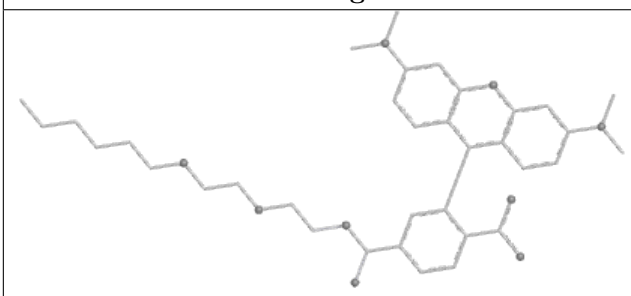
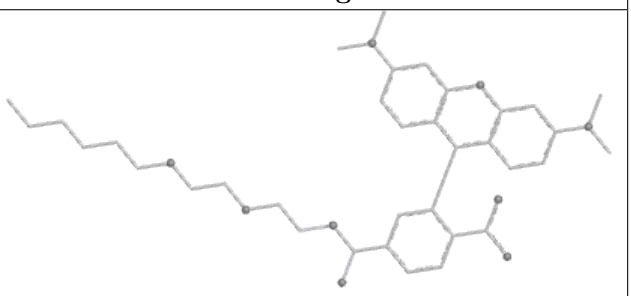


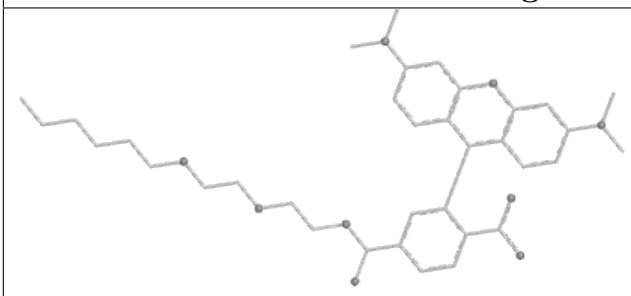
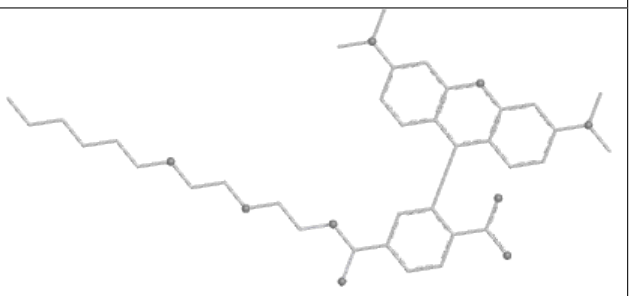
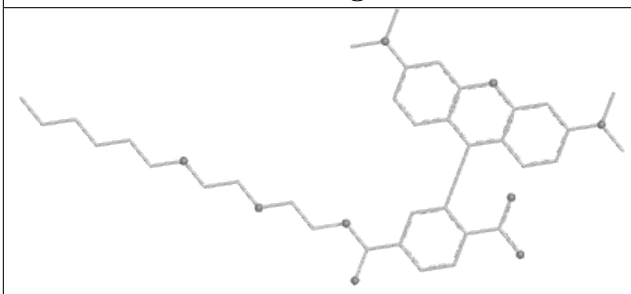
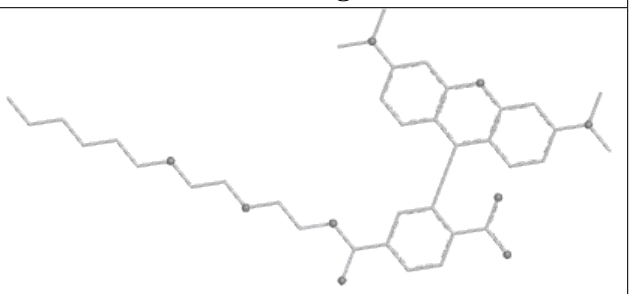
Torsions

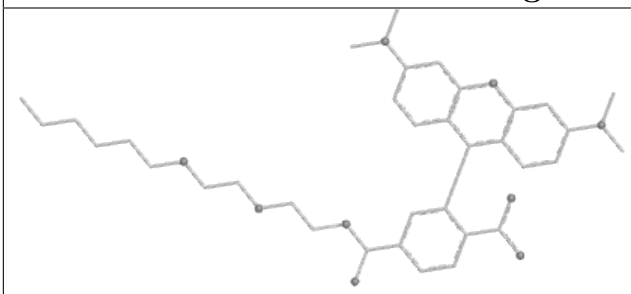
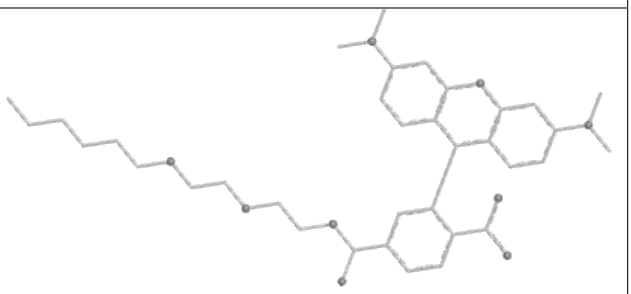
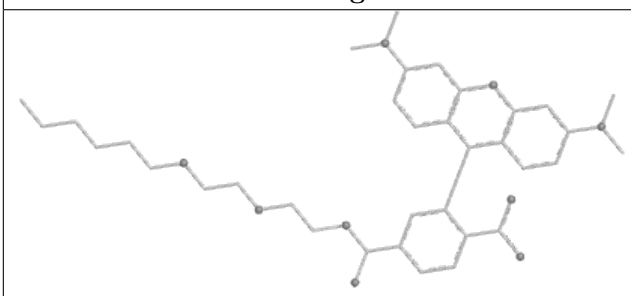
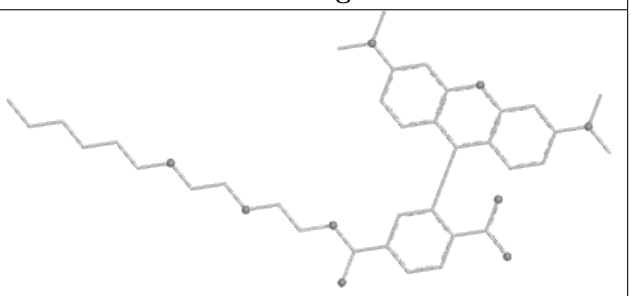


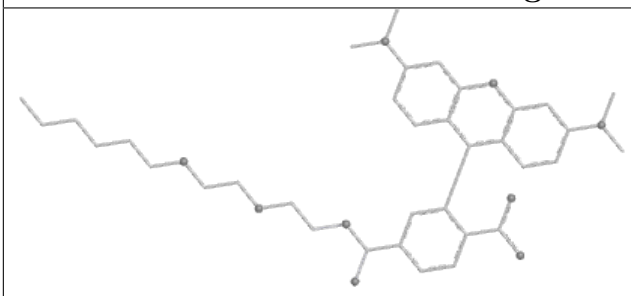
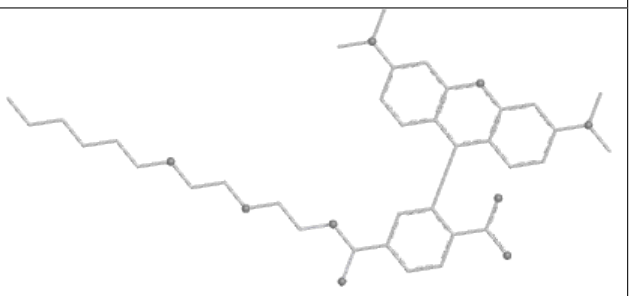
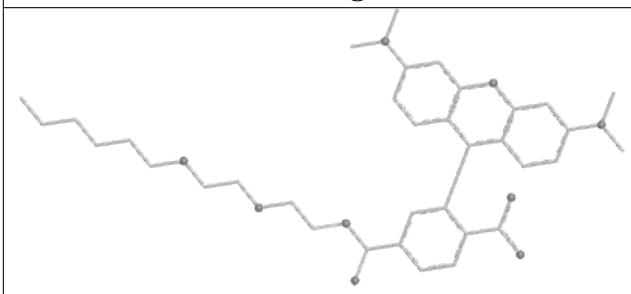
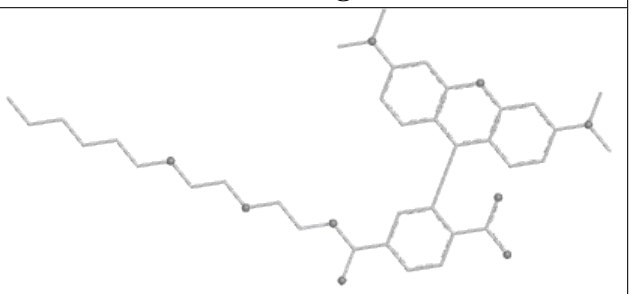
Rings

Ligand OEH C 301 (A)	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand OEH C 301 (B)	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand OEH D 301	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand OEH E 301	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand OEH F 301	
 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	293/293 (100%)	-0.43	4 (1%)	75	74	7, 12, 22, 38	0
1	B	293/293 (100%)	-0.46	4 (1%)	75	74	6, 11, 22, 35	0
1	C	293/293 (100%)	-0.38	2 (0%)	87	86	7, 12, 21, 40	0
1	D	293/293 (100%)	-0.44	3 (1%)	82	80	7, 12, 22, 36	0
1	E	293/293 (100%)	-0.38	3 (1%)	82	80	7, 12, 22, 39	0
1	F	293/293 (100%)	-0.43	3 (1%)	82	80	7, 12, 22, 35	0
All	All	1758/1758 (100%)	-0.42	19 (1%)	80	79	6, 12, 22, 40	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	3	GLY	5.4
1	D	3	GLY	4.9
1	E	3	GLY	3.9
1	A	3	GLY	3.9
1	F	3	GLY	3.7

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)

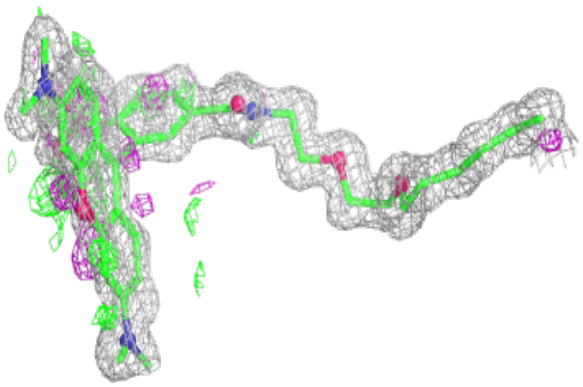
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

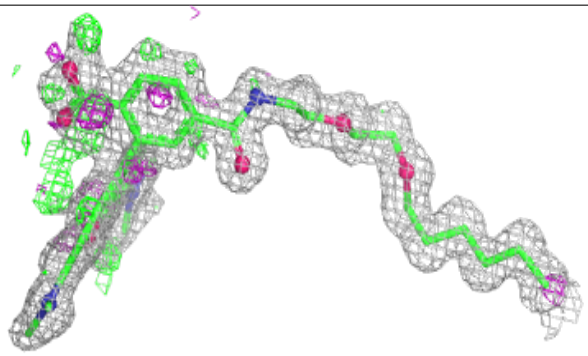
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	GOL	B	303	6/6	0.82	0.17	19,23,26,29	0
4	GOL	E	303	6/6	0.82	0.17	20,23,25,25	0
4	GOL	D	303	6/6	0.84	0.18	19,23,24,24	0
4	GOL	C	303	6/6	0.87	0.15	20,23,23,27	0
4	GOL	F	303	6/6	0.90	0.10	17,20,21,22	0
2	OEH	D	301	44/45	0.91	0.13	7,19,33,46	0
2	OEH	F	301	44/45	0.93	0.10	9,13,21,33	0
4	GOL	A	303	6/6	0.93	0.13	18,18,20,20	0
2	OEH	B	301	44/45	0.93	0.10	8,16,23,27	0
2	OEH	C	301[A]	44/45	0.93	0.15	8,14,19,24	44
2	OEH	C	301[B]	44/45	0.93	0.15	8,15,21,24	44
2	OEH	A	301	44/45	0.93	0.11	9,14,22,37	0
2	OEH	E	301	44/45	0.93	0.11	8,17,26,36	0
4	GOL	E	304	6/6	0.97	0.09	16,17,17,21	0
3	CL	C	302	1/1	0.99	0.04	8,8,8,8	0
3	CL	D	302	1/1	1.00	0.03	8,8,8,8	0
3	CL	E	302	1/1	1.00	0.02	7,7,7,7	0
3	CL	F	302	1/1	1.00	0.03	8,8,8,8	0
3	CL	B	302	1/1	1.00	0.02	7,7,7,7	0
3	CL	A	302	1/1	1.00	0.05	8,8,8,8	0

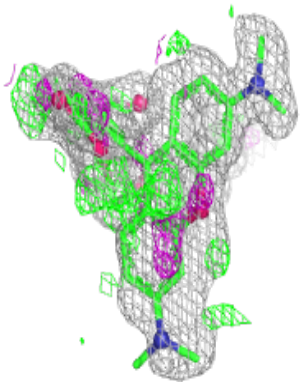
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around OEH D 301:

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

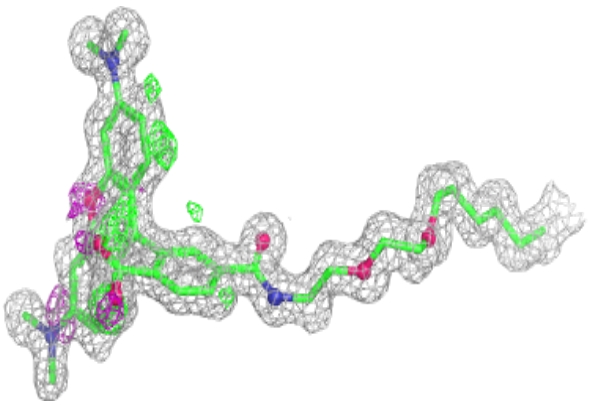


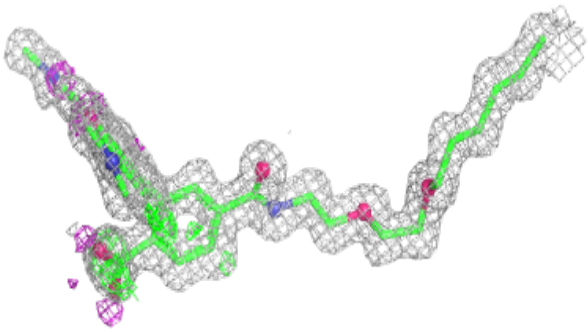


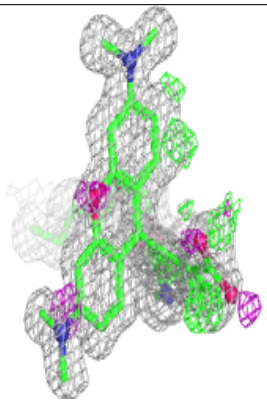


Electron density around OEH F 301:

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

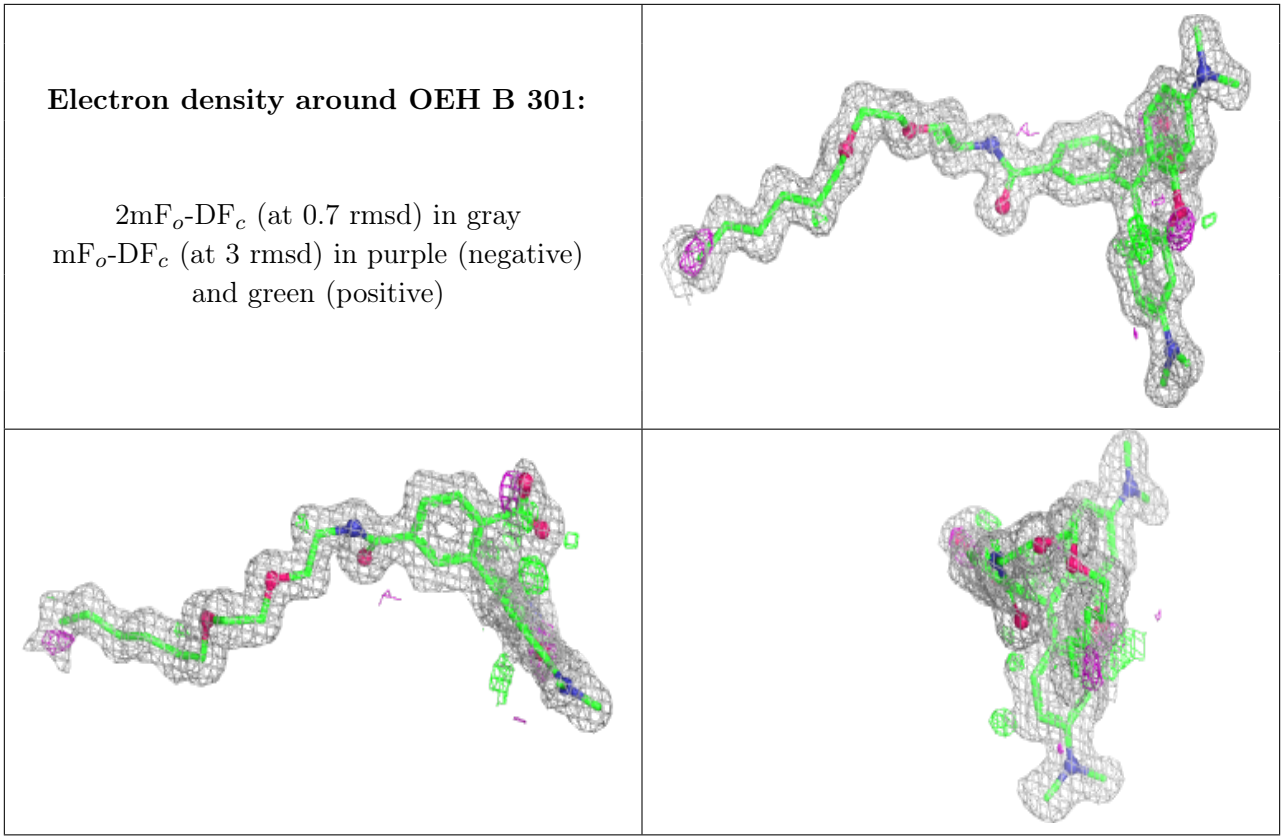






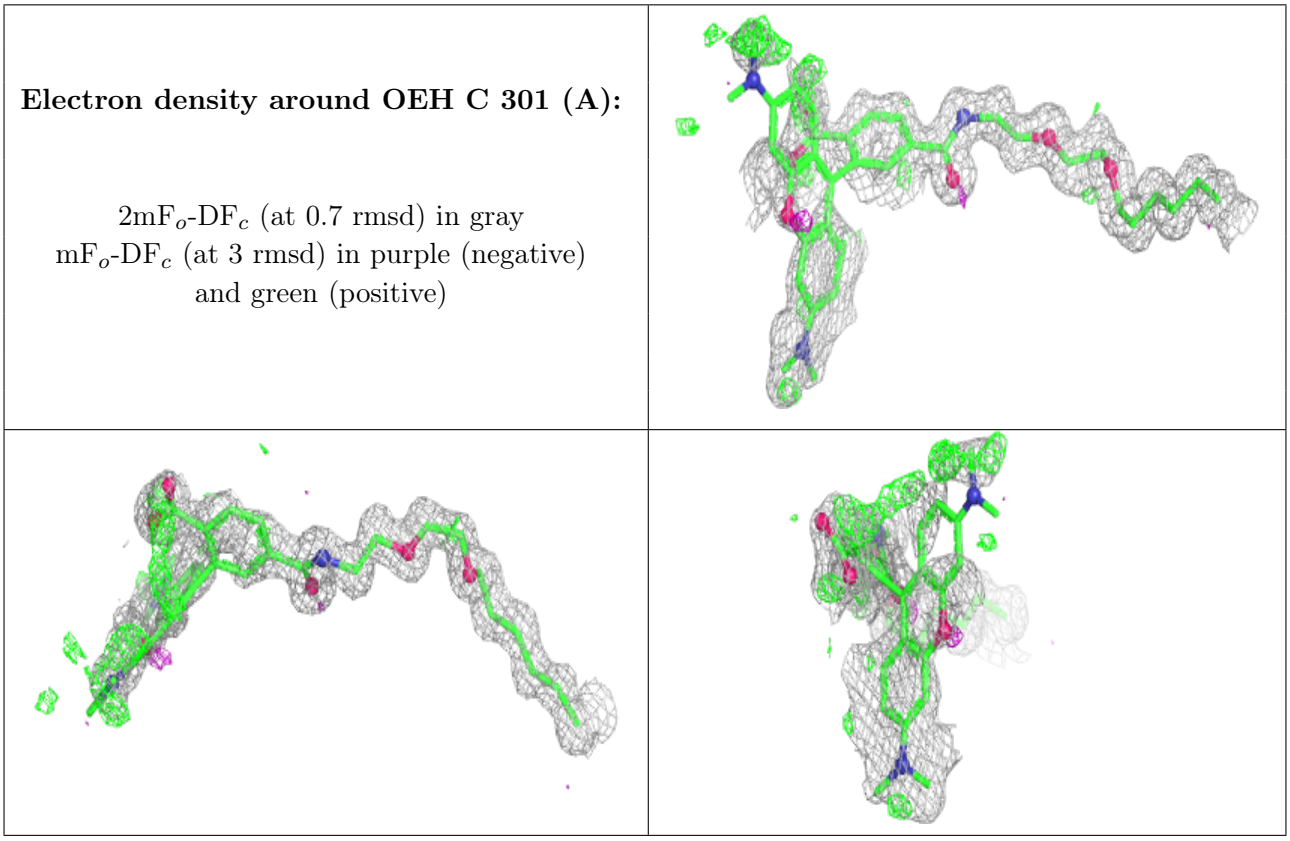
Electron density around OEH B 301:

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



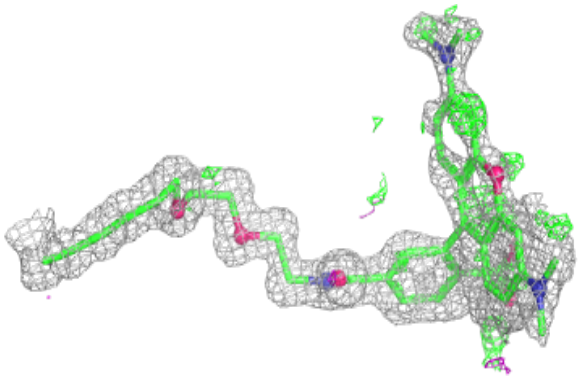
Electron density around OEH C 301 (A):

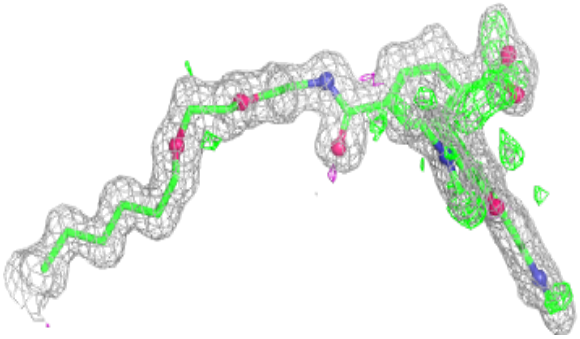
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

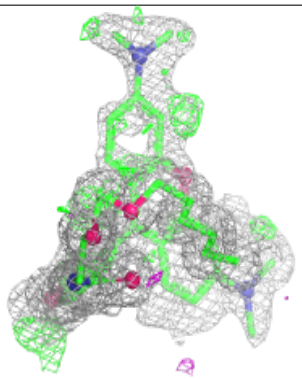


Electron density around OEH C 301 (B):

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

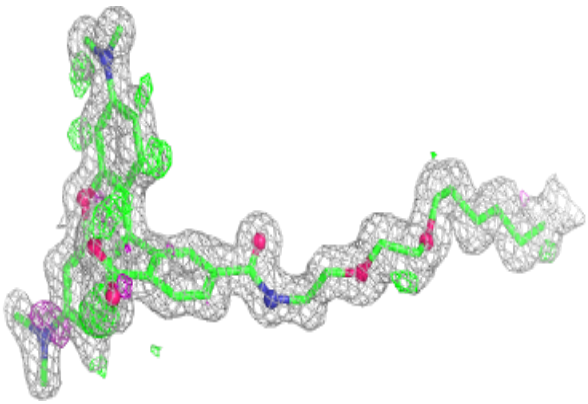


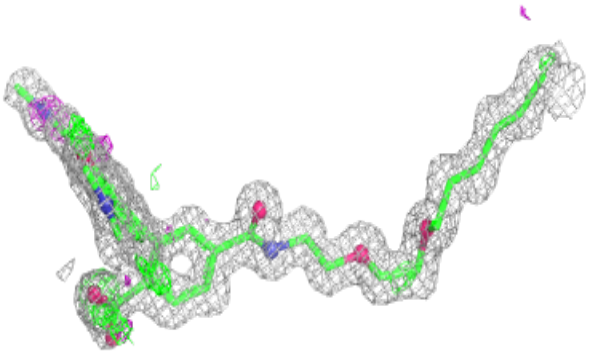


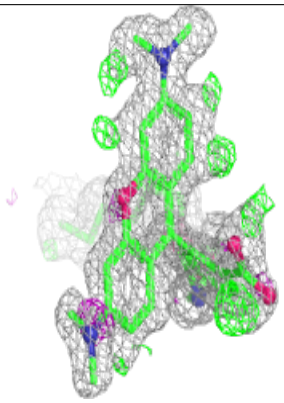


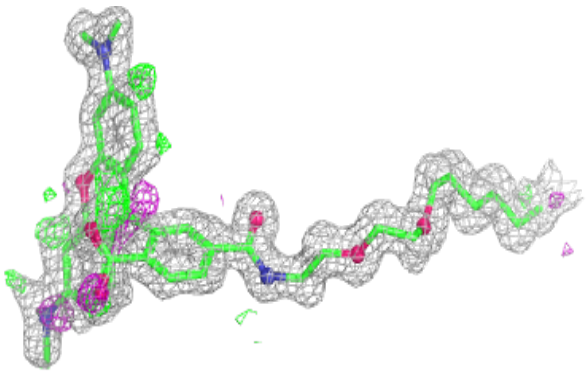
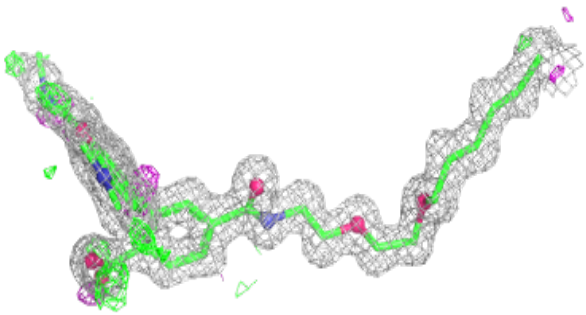
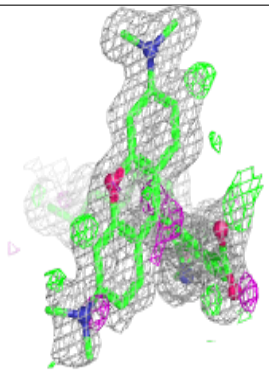
Electron density around OEH A 301:

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







Electron density around OEH E 301: 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.