wwPDB X-ray Structure Validation Summary Report (i)

Sep 16, 2021 – 02:15 pm BST

PDB ID	:	7OPN
Title	:	Human Aldehyde Oxidase SNP R1231H in complex with Raloxifene
Authors	:	Mota, C.; Coelho, C.; Santos Silva, T.; Romao, M.J.
Deposited on	:	2021-06-01
Resolution	:	2.60 Å(reported)

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

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

The following versions of software and data (see references (1)) were used in the production of this report:

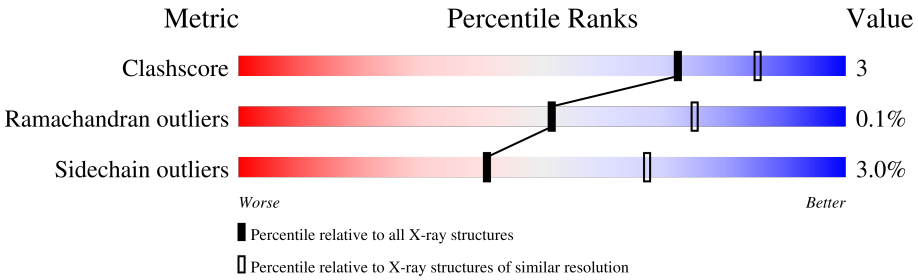
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.1

1 Overall quality at a glance

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

The reported resolution of this entry is 2.60 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))
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	1338	86% 10% .
1	B	1338	86% 11% .

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
10	DMS	B	3008	-	X	-	-

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 20650 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.

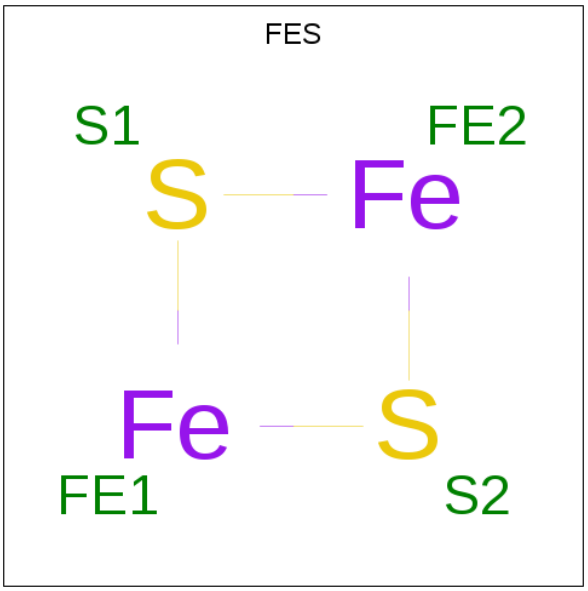
- Molecule 1 is a protein called Aldehyde oxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1295	Total	C	N	O	S	0	0	0
			10055	6389	1731	1855	80			
1	B	1295	Total	C	N	O	S	0	1	0
			10064	6397	1731	1856	80			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1231	HIS	ARG	engineered mutation	UNP Q06278
B	1231	HIS	ARG	engineered mutation	UNP Q06278

- Molecule 2 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe_2S_2).



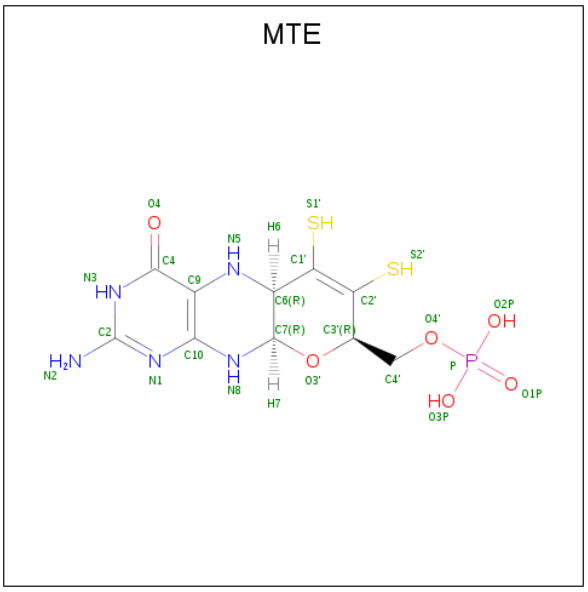
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	Fe	S	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

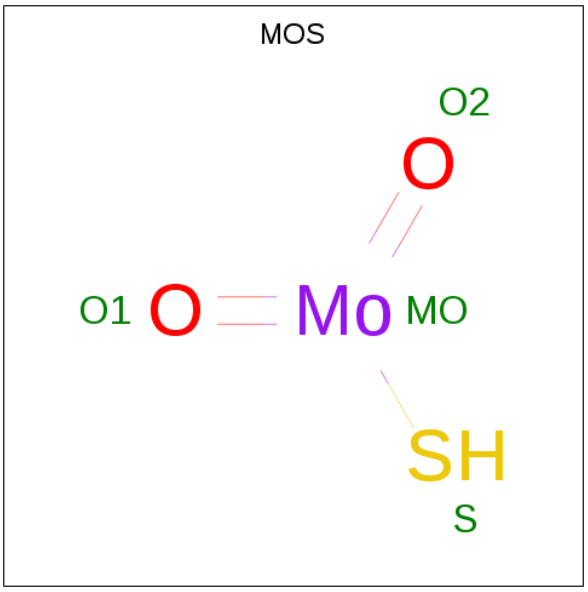
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Fe S 4 2 2	0	0
2	B	1	Total Fe S 4 2 2	0	0
2	B	1	Total Fe S 4 2 2	0	0

- Molecule 3 is PHOSPHONIC ACIDMONO-(2-AMINO-5,6-DIMERCAPTO-4-OXO-3,7,8A, 9,10,10A-HEXAHYDRO-4H-8-OXA-1,3,9,10-TETRAAZA-ANTHRACEN-7-YLMETHYL) ESTER (three-letter code: MTE) (formula: $C_{10}H_{14}N_5O_6PS_2$).



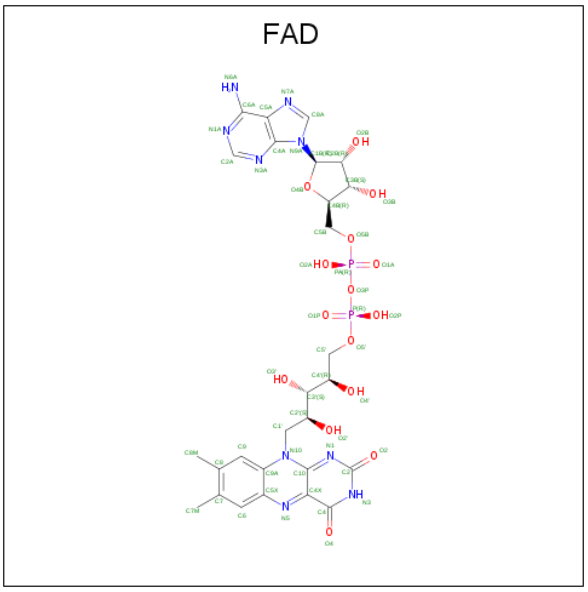
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O P S 24 10 5 6 1 2	0	0
3	B	1	Total C N O P S 24 10 5 6 1 2	0	0

- Molecule 4 is DIOXOTHIOMOLYBDENUM(VI) ION (three-letter code: MOS) (formula: $HMoO_2S$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	Mo	O	S	0	0
			4	1	2	1		
4	B	1	Total	Mo	O	S	0	0
			4	1	2	1		

- Molecule 5 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).



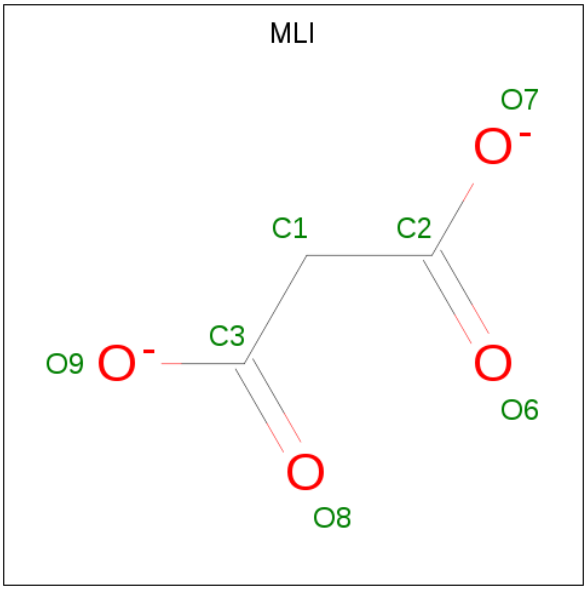
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

Continued on next page...

Continued from previous page...

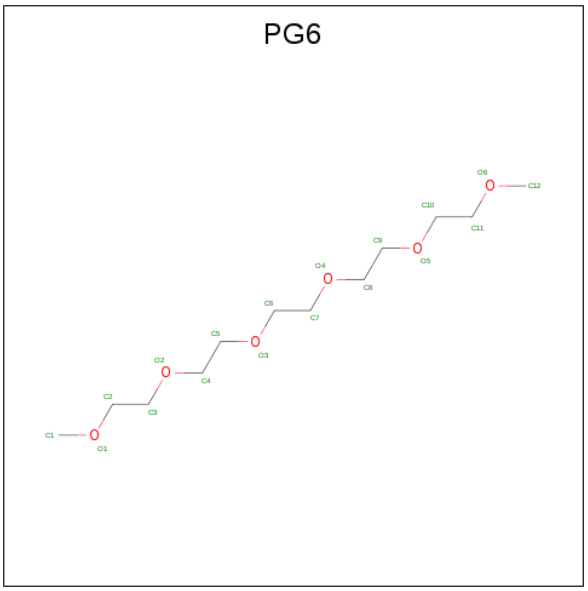
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 6 is MALONATE ION (three-letter code: MLI) (formula: $C_3H_2O_4$).



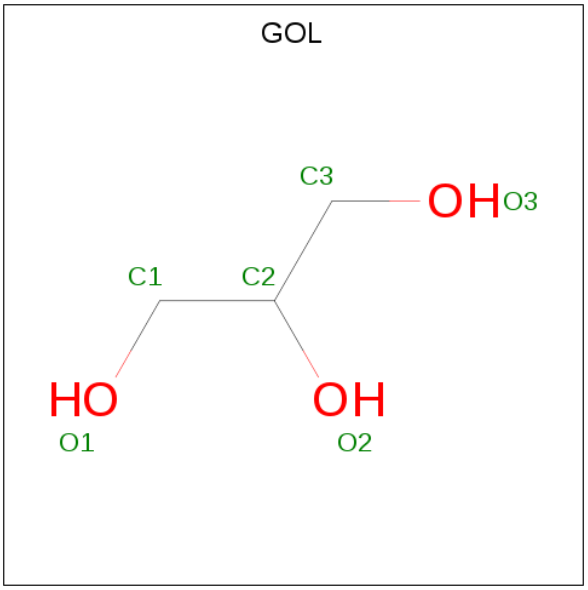
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			7	3	4		
6	B	1	Total	C	O	0	0
			7	3	4		

- Molecule 7 is 1-(2-METHOXY-ETHOXY)-2-{2-[2-(2-METHOXY-ETHOXY]-ETHOXY}-ETHANE (three-letter code: PG6) (formula: $C_{12}H_{26}O_6$).



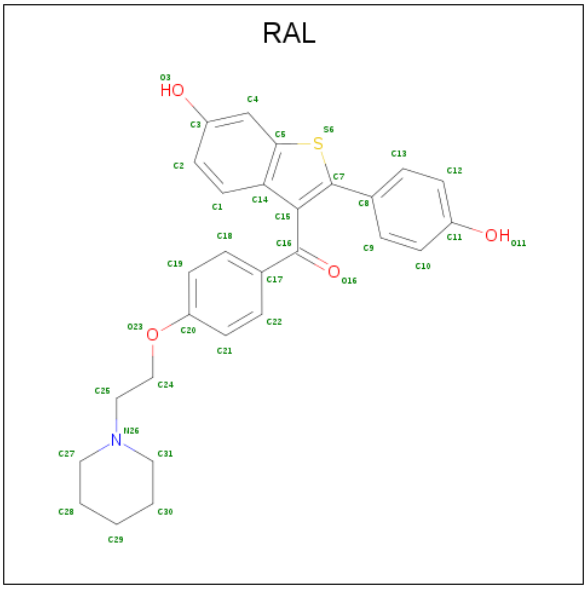
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 18 12 6	0	0

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



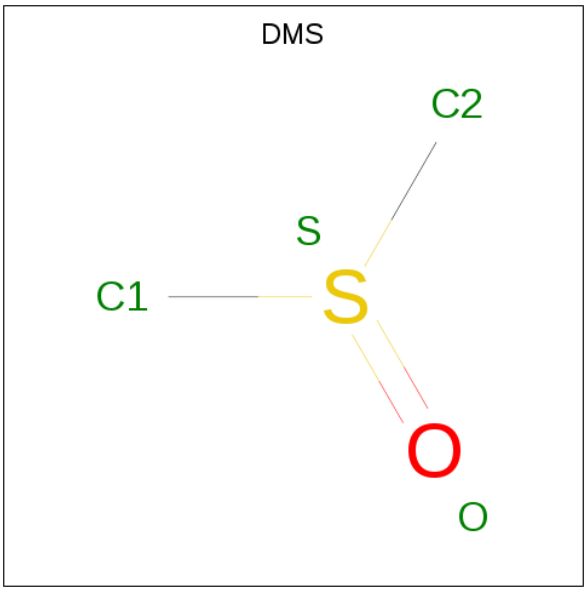
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C O 6 3 3	0	0
8	A	1	Total C O 6 3 3	0	0
8	B	1	Total C O 6 3 3	0	0

- Molecule 9 is RALOXIFENE (three-letter code: RAL) (formula: $C_{28}H_{27}NO_4S$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	A	1	Total	C	N	O	S	0	0
			34	28	1	4	1		
9	A	1	Total	C	N	O	S	0	0
			34	28	1	4	1		
9	B	1	Total	C	N	O	S	0	0
			34	28	1	4	1		
9	B	1	Total	C	N	O	S	0	0
			34	28	1	4	1		

- Molecule 10 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).



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

- Molecule 11 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	91	Total O 91 91	0	0
11	B	72	Total O 72 72	0	0

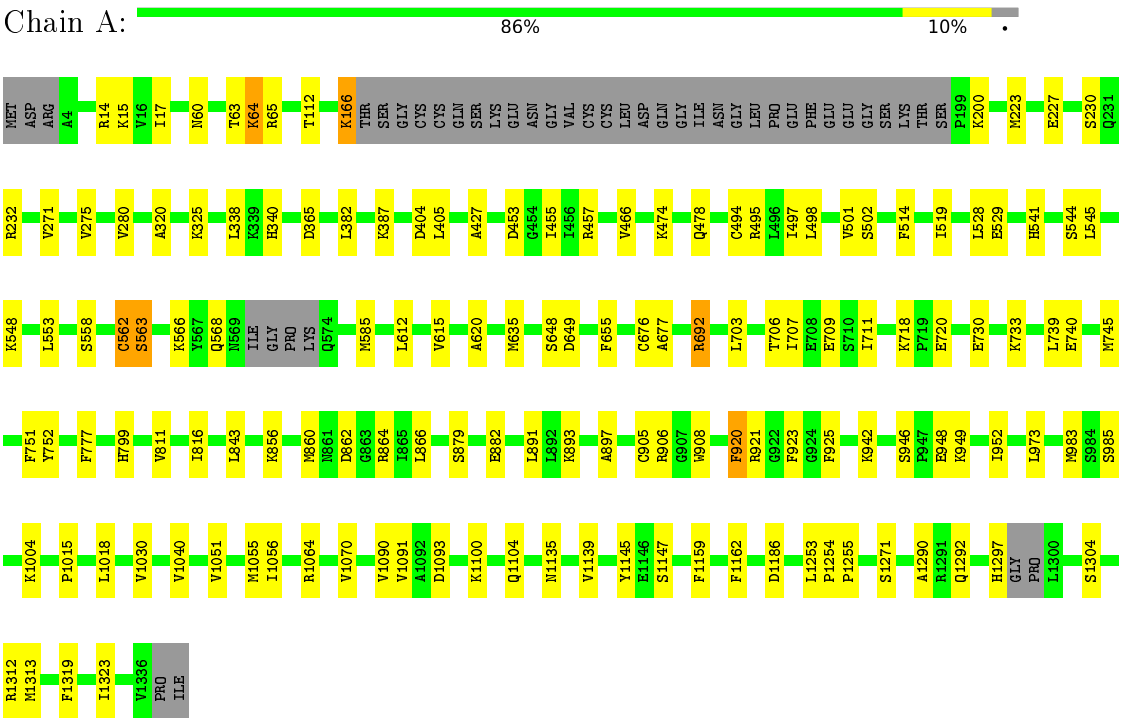
3 Residue-property plots (i)

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. 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. 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.

Note EDS was not executed.

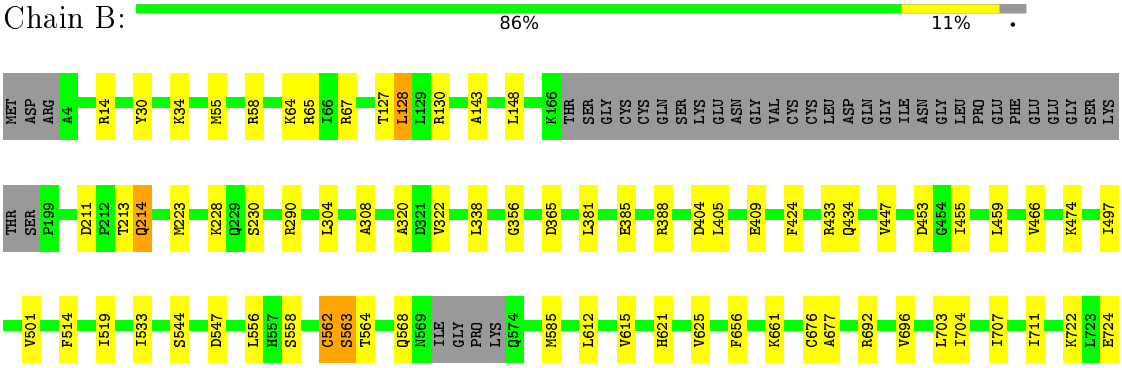
- Molecule 1: Aldehyde oxidase

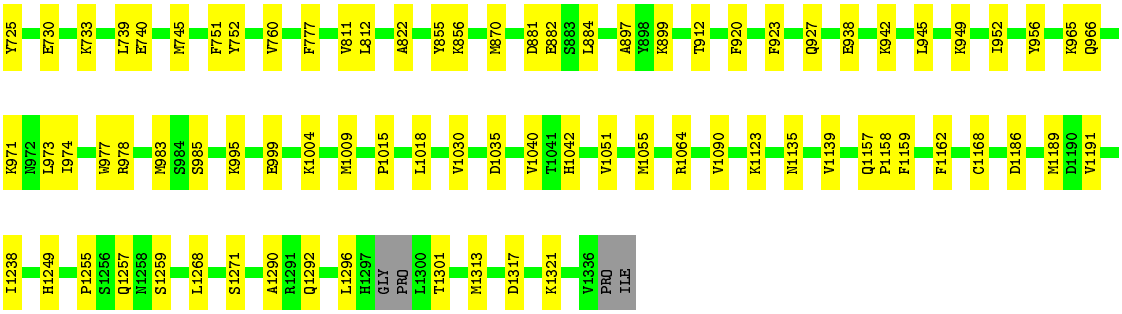
Chain A:



- Molecule 1: Aldehyde oxidase

Chain B:




Y725
E730
K733
L739
E740
M745
F751
Y752
V760
F777
V811
L812
A822
Y855
K856
M870
D881
E882
S883
L884
A897
Y898
K899
T912
F920
F923
Q927
E938
K942
L945
K949
I952
Y956
K965
Q966
K971
N972
L973
I974
W977
R978
M983
S984
S985
K995
E999
K1004
M1009
P1015
L1018
V1030
D1035
V1040
T1041
H1042
V1051
M1055
R1064
V1090
K1123
N1135
V1139
Q1157
P1158
F1159
F1162
C1168
D1186
M1189
D1190
V1191
I1238
H1249
P1255
S1256
Q1257
N1258
S1259
L1268
S1271
A1290
R1291
Q1292
L1296
H1297
GLY
PRO
L1300
T1301
M1313
D1317
K1321
V1336
PRO
ILE

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	149.53Å 149.53Å 269.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.01 – 2.60	Depositor
% Data completeness (in resolution range)	99.9 (49.01-2.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.206 , 0.241	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	20650	wwPDB-VP
Average B, all atoms (Å^2)	60.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MLI, MTE, GOL, FES, MOS, FAD, PG6, RAL, DMS

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	0.25	0/10269	0.48	0/13887
1	B	0.25	0/10282	0.48	0/13905
All	All	0.25	0/20551	0.48	0/27792

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	10055	0	10093	69	0
1	B	10064	0	10102	70	0
2	A	8	0	0	0	0
2	B	8	0	0	0	0
3	A	24	0	10	0	0
3	B	24	0	10	0	0
4	A	4	0	0	0	0
4	B	4	0	0	0	0
5	A	53	0	31	0	0
5	B	53	0	31	0	0
6	A	7	0	2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	7	0	2	0	0
7	A	18	0	26	2	0
8	A	12	0	16	0	0
8	B	6	0	8	1	0
9	A	68	0	52	0	0
9	B	68	0	52	1	0
10	B	4	0	6	0	0
11	A	91	0	0	0	0
11	B	72	0	0	0	0
All	All	20650	0	20441	139	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

The worst 5 of 139 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:995:LYS:NZ	1:B:999:GLU:OE2	2.24	0.70
1:B:564:THR:OG1	1:B:1249:HIS:ND1	2.27	0.67
1:A:862:ASP:OD2	1:A:864:ARG:NH1	2.29	0.65
1:B:945:LEU:HD22	1:B:949:LYS:HE2	1.78	0.65
1:A:200:LYS:H	1:A:200:LYS:HD2	1.62	0.64

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	1287/1338 (96%)	1256 (98%)	30 (2%)	1 (0%)	51 75
1	B	1288/1338 (96%)	1258 (98%)	29 (2%)	1 (0%)	51 75
All	All	2575/2676 (96%)	2514 (98%)	59 (2%)	2 (0%)	51 75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	811	VAL
1	B	811	VAL

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	1100/1136 (97%)	1067 (97%)	33 (3%)	41	67
1	B	1101/1136 (97%)	1068 (97%)	33 (3%)	41	67
All	All	2201/2272 (97%)	2135 (97%)	66 (3%)	41	67

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	923	PHE
1	B	985	SER
1	B	1301	THR
1	A	882	GLU
1	A	879	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	69	HIS
1	B	214	GLN
1	B	298	ASN
1	B	415	ASN
1	B	1292	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

21 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	MTE	A	3003	4	21,26,26	0.98	1 (4%)	21,40,40	2.03	7 (33%)
2	FES	A	3001	1	0,4,4	0.00	-	-		
8	GOL	A	3008	-	5,5,5	0.90	0	5,5,5	0.99	0
4	MOS	B	3004	3	0,3,3	0.00	-	-		
9	RAL	A	3011	-	33,38,38	0.60	1 (3%)	43,53,53	0.88	2 (4%)
6	MLI	A	3006	-	0,6,6	0.00	-	0,7,7	0.00	-
5	FAD	B	3005	-	51,58,58	1.08	2 (3%)	60,89,89	1.73	7 (11%)
2	FES	B	3002	1	0,4,4	0.00	-	-		
10	DMS	B	3008	-	3,3,3	0.73	0	3,3,3	3.24	3 (100%)
7	PG6	A	3007	-	17,17,17	0.53	0	16,16,16	0.18	0
4	MOS	A	3004	3	0,3,3	0.00	-	-		
2	FES	A	3002	1	0,4,4	0.00	-	-		
9	RAL	B	3010	-	33,38,38	0.54	1 (3%)	43,53,53	0.96	2 (4%)
8	GOL	B	3006	-	5,5,5	0.92	0	5,5,5	1.01	0
6	MLI	B	3007	-	0,6,6	0.00	-	0,7,7	0.00	-
9	RAL	B	3009	-	33,38,38	0.63	1 (3%)	43,53,53	0.83	2 (4%)
3	MTE	B	3003	4	21,26,26	0.99	1 (4%)	21,40,40	2.00	7 (33%)
2	FES	B	3001	1	0,4,4	0.00	-	-		
5	FAD	A	3005	-	51,58,58	1.07	2 (3%)	60,89,89	1.72	7 (11%)
8	GOL	A	3009	-	5,5,5	0.92	0	5,5,5	0.98	0
9	RAL	A	3010	-	33,38,38	0.60	1 (3%)	43,53,53	0.86	2 (4%)

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	FAD	B	3005	-	-	1/30/50/50	0/6/6/6
9	RAL	B	3009	-	-	2/15/26/26	0/5/5/5
3	MTE	B	3003	4	-	1/6/34/34	0/3/3/3
2	FES	B	3002	1	-	-	0/1/1/1
2	FES	B	3001	1	-	-	0/1/1/1
3	MTE	A	3003	4	-	1/6/34/34	0/3/3/3
7	PG6	A	3007	-	-	9/15/15/15	-
9	RAL	A	3010	-	-	0/15/26/26	0/5/5/5
8	GOL	A	3008	-	-	2/4/4/4	-
2	FES	A	3001	1	-	-	0/1/1/1
9	RAL	A	3011	-	-	0/15/26/26	0/5/5/5
2	FES	A	3002	1	-	-	0/1/1/1
6	MLI	A	3006	-	-	0/0/4/4	-
9	RAL	B	3010	-	-	1/15/26/26	0/5/5/5
8	GOL	B	3006	-	-	2/4/4/4	-
5	FAD	A	3005	-	-	1/30/50/50	0/6/6/6
6	MLI	B	3007	-	-	0/0/4/4	-
8	GOL	A	3009	-	-	1/4/4/4	-

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	3005	FAD	C4X-C10	5.74	1.44	1.38
5	A	3005	FAD	C4X-C10	5.62	1.44	1.38
3	B	3003	MTE	C4-N3	3.41	1.39	1.33
3	A	3003	MTE	C4-N3	3.38	1.38	1.33
9	B	3009	RAL	C15-C16	-3.08	1.47	1.50

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	3005	FAD	C4-N3-C2	8.06	121.94	115.14
5	A	3005	FAD	C4-N3-C2	8.04	121.93	115.14
5	A	3005	FAD	C10-C4X-N5	5.23	124.87	121.26
3	A	3003	MTE	C4-C9-C10	5.22	119.21	114.57
5	B	3005	FAD	C4-C4X-C10	-5.20	116.51	119.95

There are no chirality outliers.

5 of 21 torsion outliers are listed below:

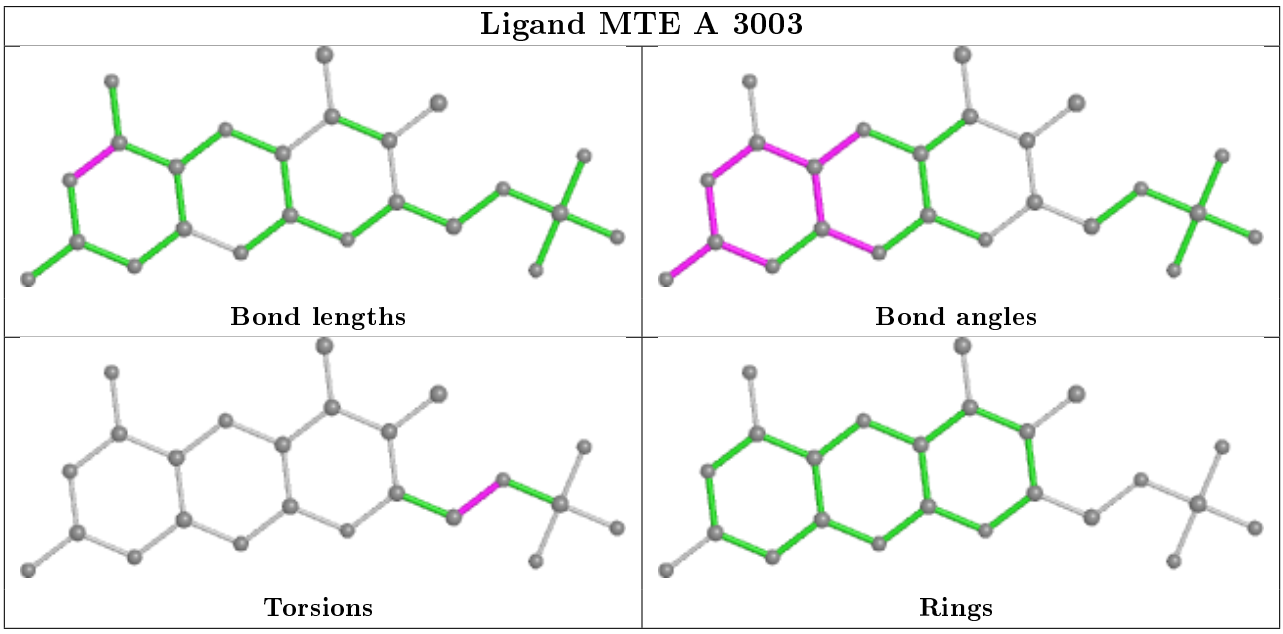
Mol	Chain	Res	Type	Atoms
3	A	3003	MTE	C3'-C4'-O4'-P
9	B	3009	RAL	O23-C24-C25-N26
7	A	3007	PG6	O1-C2-C3-O2
7	A	3007	PG6	O5-C10-C11-O6
7	A	3007	PG6	C2-C3-O2-C4

There are no ring outliers.

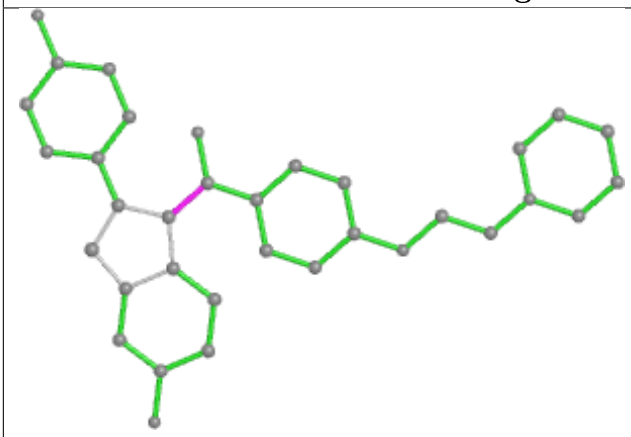
3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	3007	PG6	2	0
8	B	3006	GOL	1	0
9	B	3009	RAL	1	0

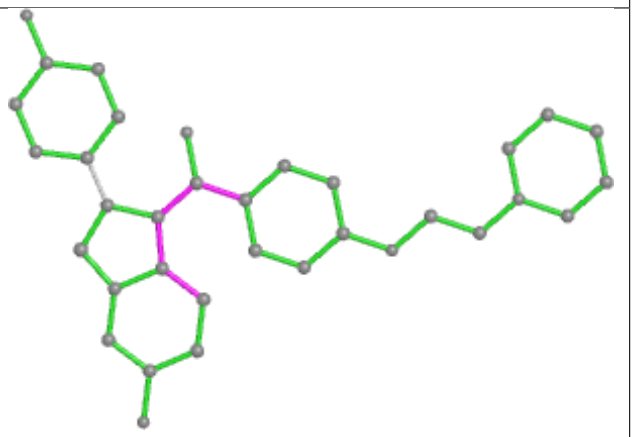
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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



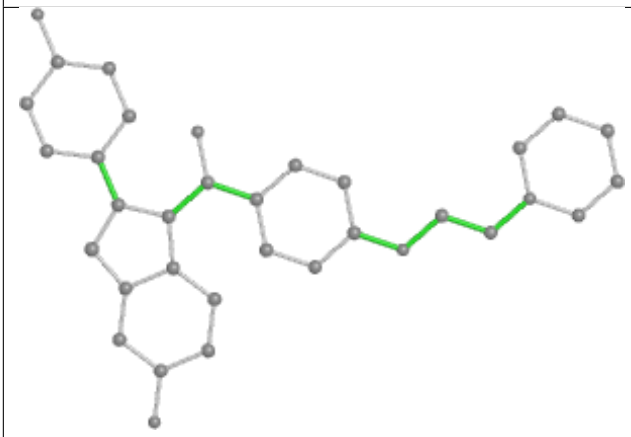
Ligand RAL A 3011



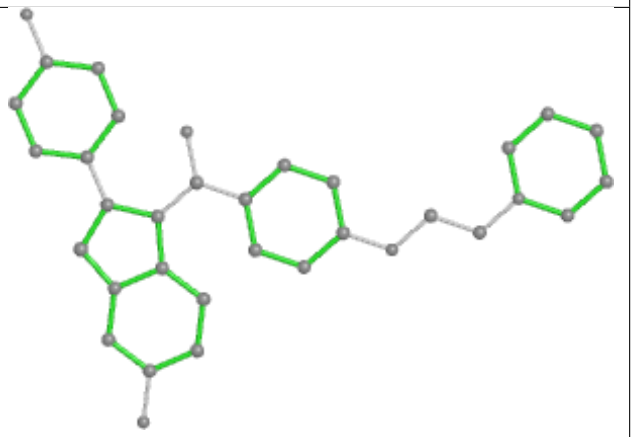
Bond lengths



Bond angles

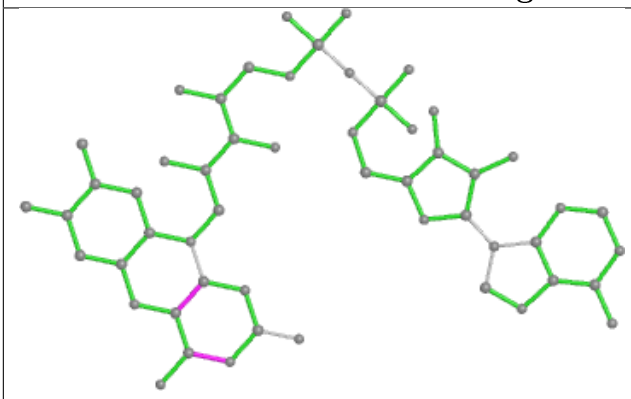


Torsions

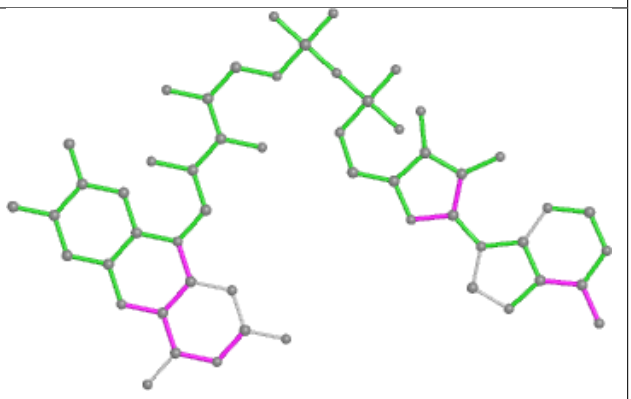


Rings

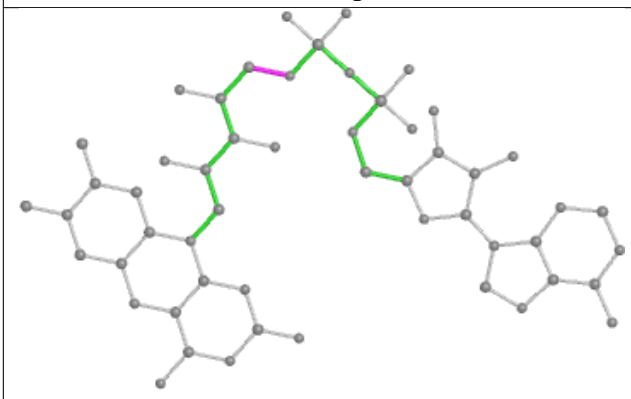
Ligand FAD B 3005



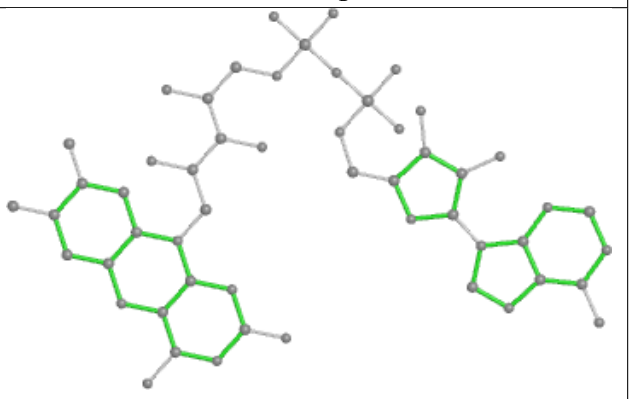
Bond lengths



Bond angles

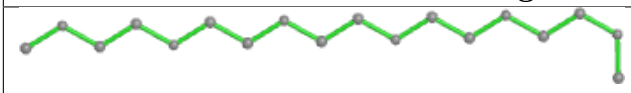


Torsions

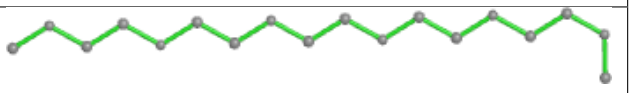


Rings

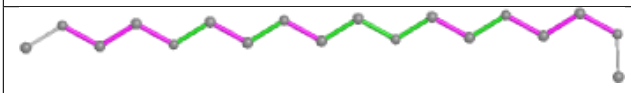
Ligand PG6 A 3007



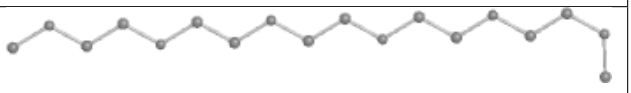
Bond lengths



Bond angles

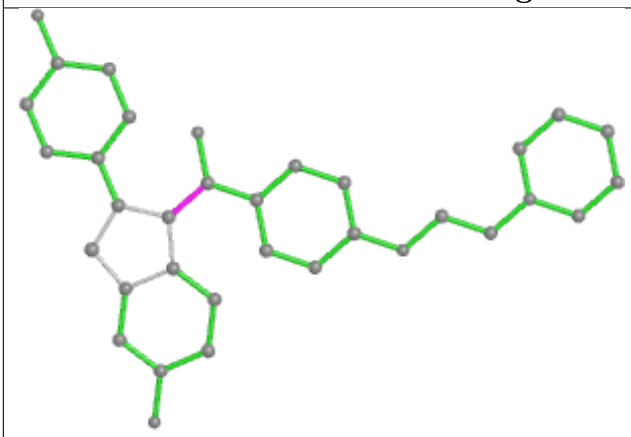


Torsions

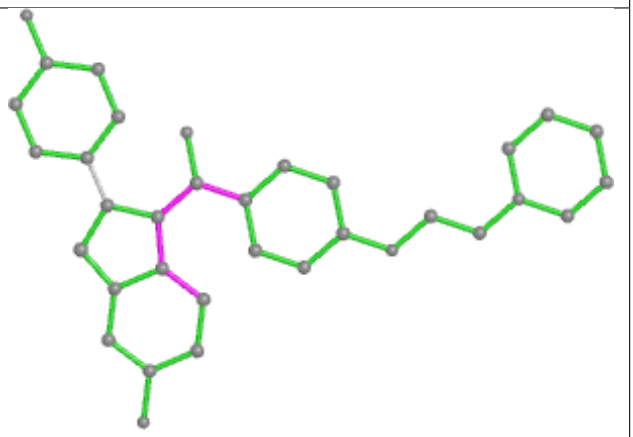


Rings

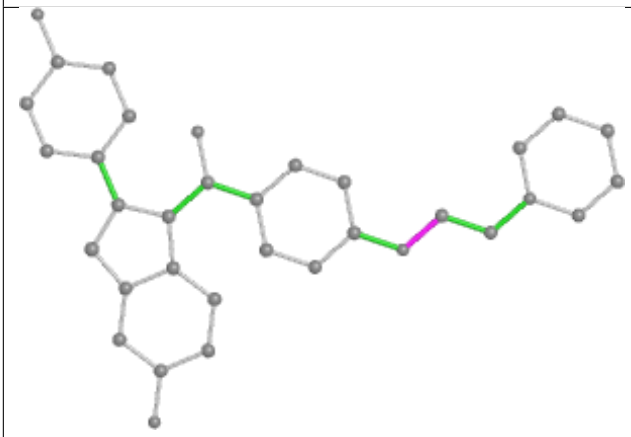
Ligand RAL B 3010



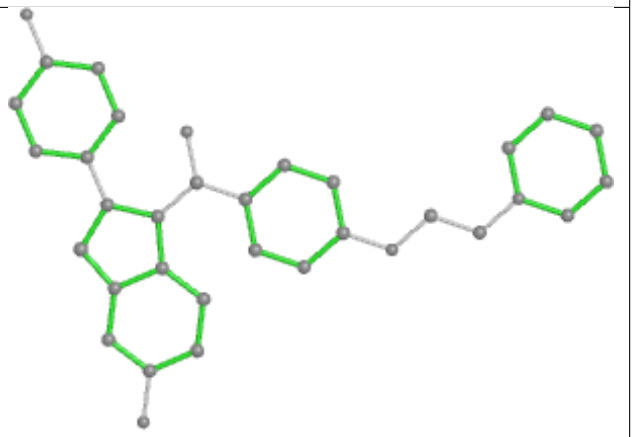
Bond lengths



Bond angles

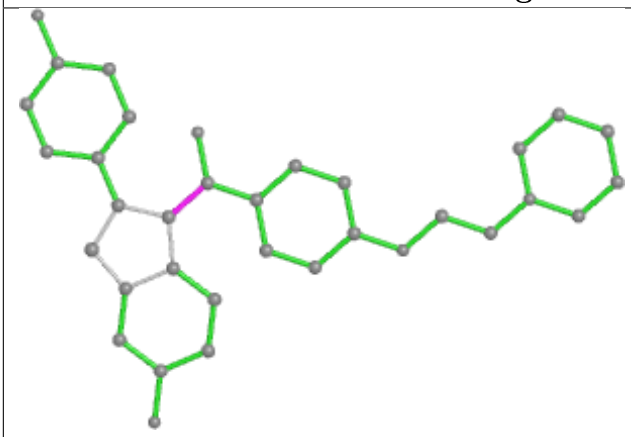


Torsions

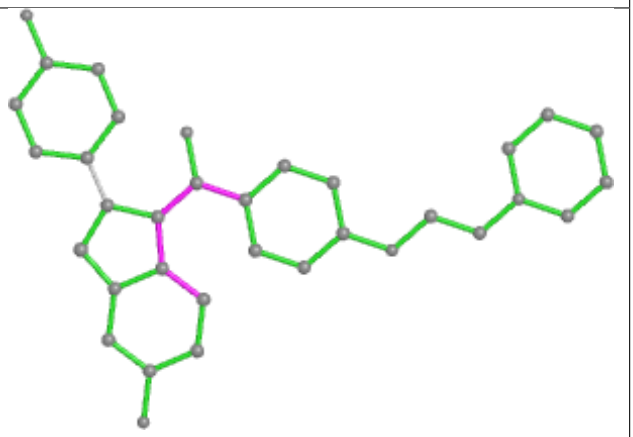


Rings

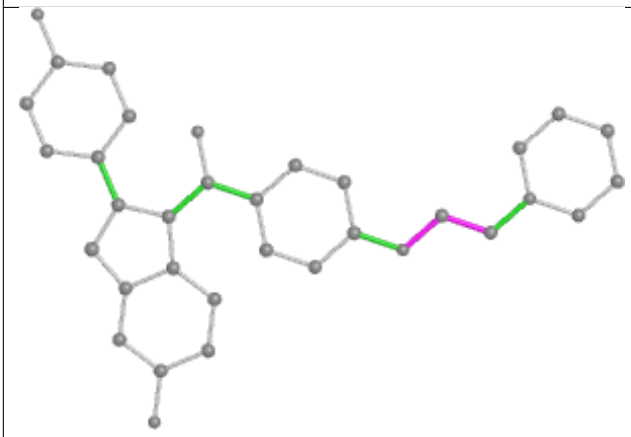
Ligand RAL B 3009



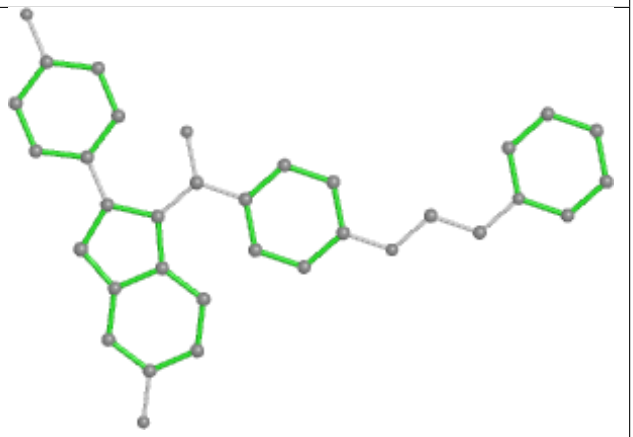
Bond lengths



Bond angles

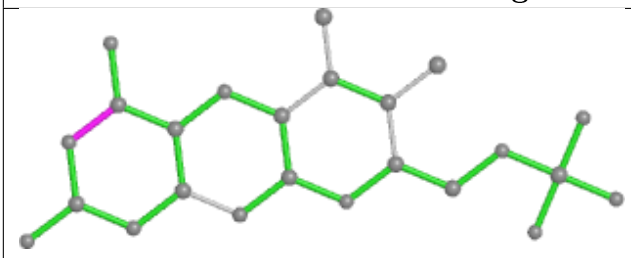


Torsions

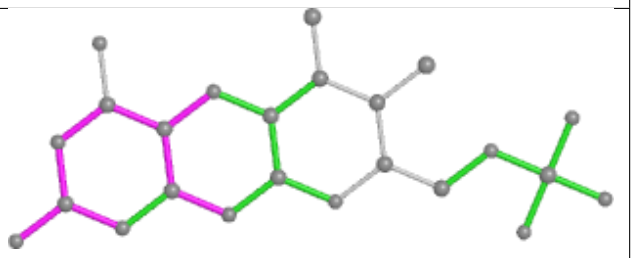


Rings

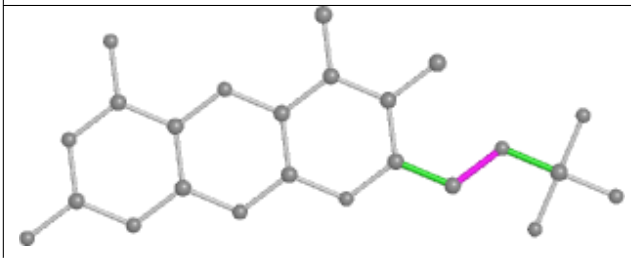
Ligand MTE B 3003



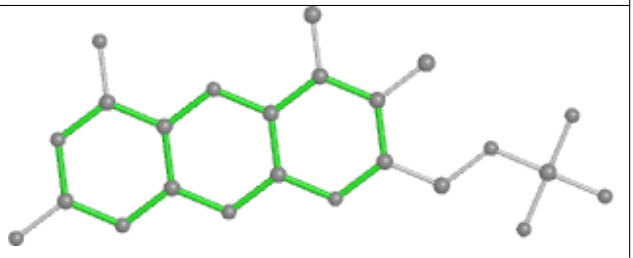
Bond lengths



Bond angles

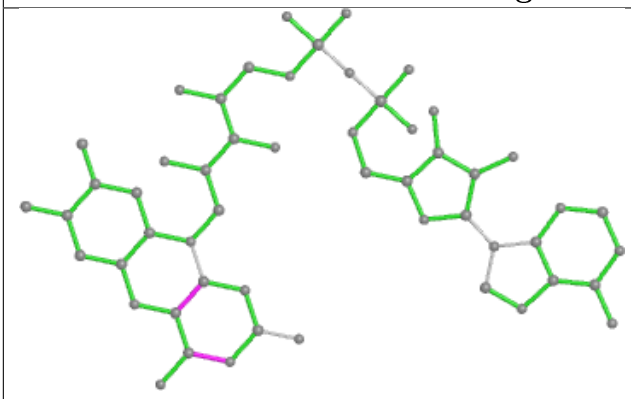


Torsions

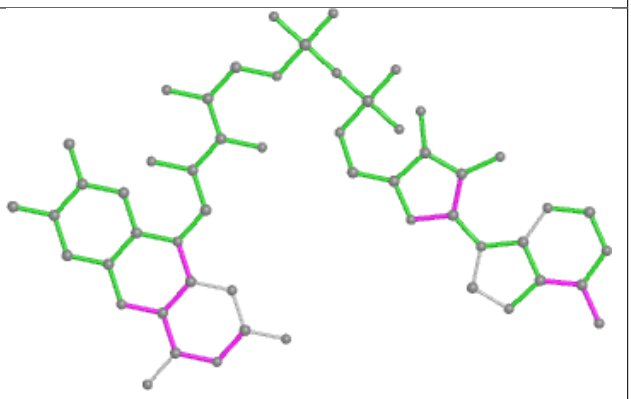


Rings

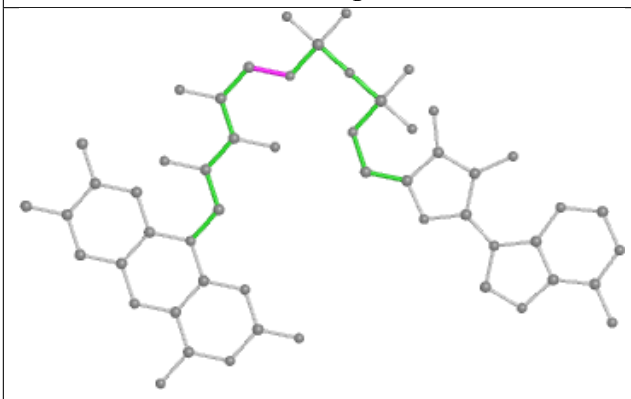
Ligand FAD A 3005



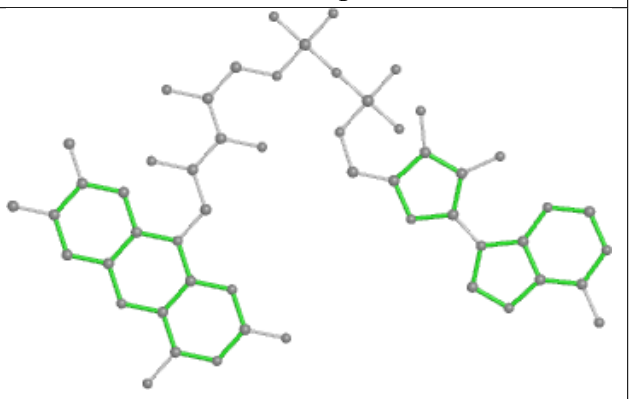
Bond lengths



Bond angles

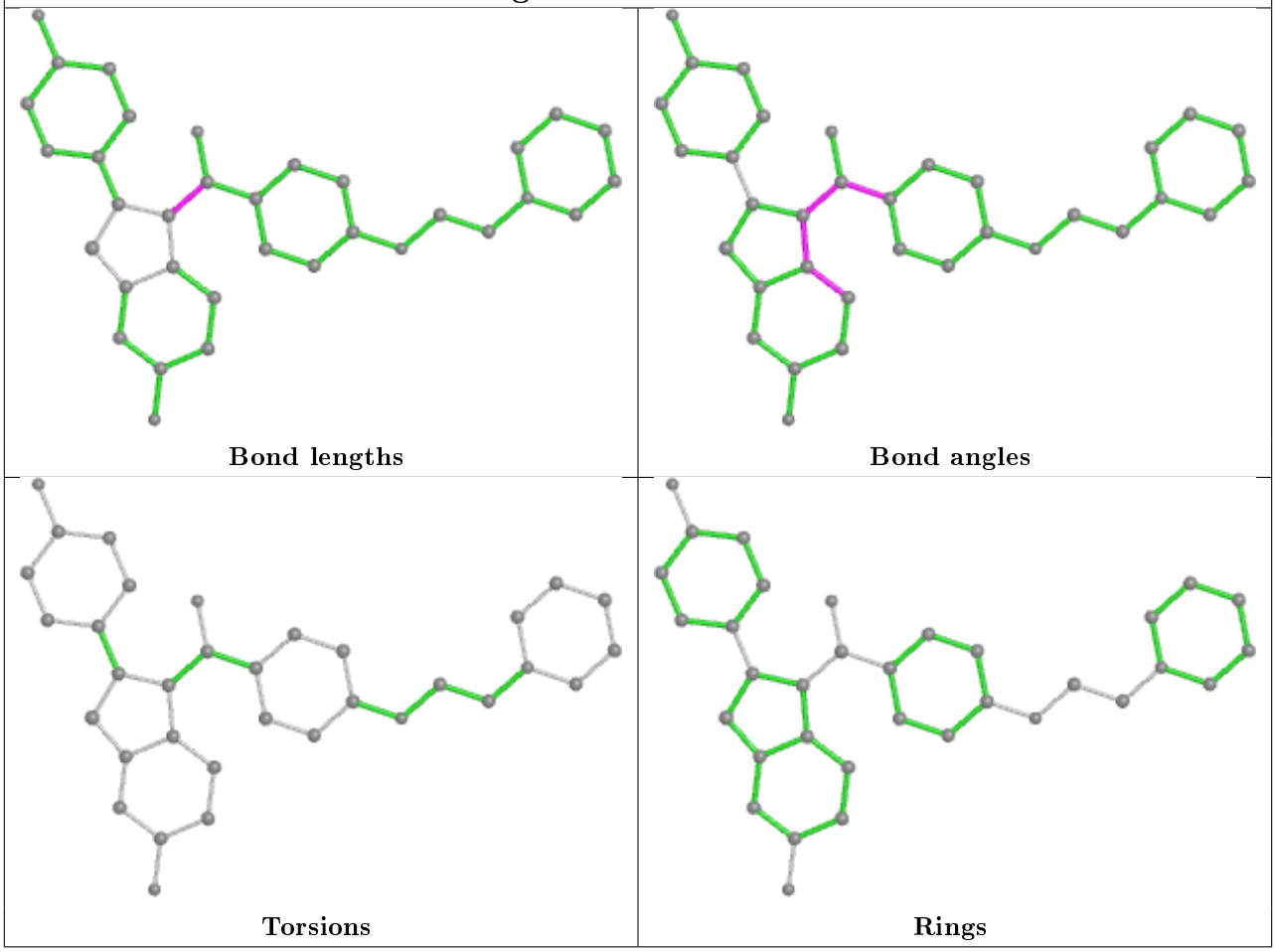


Torsions



Rings

Ligand RAL A 3010



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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.