



# wwPDB EM Validation Summary Report i

Jul 15, 2025 – 10:53 PM JST

PDB ID : 8Y01 / pdb\_00008y01  
EMDB ID : EMD-38800  
Title : Cryo-EM structure of Medium-wave-sensitive opsin 1  
Authors : Peng, Q.; Jiang, H.H.; Cheng, X.Y.; Li, J.; Zhang, J.  
Deposited on : 2024-01-21  
Resolution : 2.48 Å (reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>  
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](#) ①) were used in the production of this report:

EMDB validation analysis : **FAILED**  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
MolProbity : 4-5-2 with Phenix2.0rc1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.44

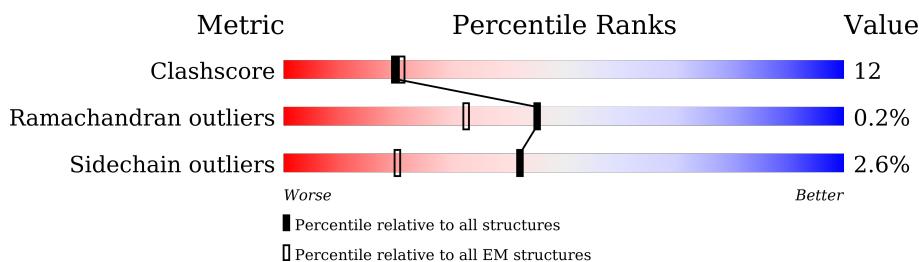
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 2.48 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)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ .



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
6	RET	R	401	-	-	X	-

## 2 Entry composition [\(i\)](#)

There are 6 unique types of molecules in this entry. The entry contains 8705 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Guanine nucleotide-binding protein G(i) subunit alpha-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	205	Total	C	N	O	S	0	0
			1570	1010	263	285	12		

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	47	CYS	SER	engineered mutation	UNP P63096
A	202	THR	GLY	engineered mutation	UNP P63096
A	203	ALA	GLY	engineered mutation	UNP P63096
A	245	ALA	GLU	engineered mutation	UNP P63096
A	326	SER	ALA	engineered mutation	UNP P63096

- Molecule 2 is a protein called Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	339	Total	C	N	O	S	0	0
			2599	1603	468	507	21		

- Molecule 3 is a protein called Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	G	58	Total	C	N	O	S	0	0
			438	274	76	85	3		

- Molecule 4 is a protein called scFv Recombinant Human Monoclonal Antibody (scFv16).

Mol	Chain	Residues	Atoms					AltConf	Trace
4	S	233	Total	C	N	O	S	0	0
			1751	1116	295	330	10		

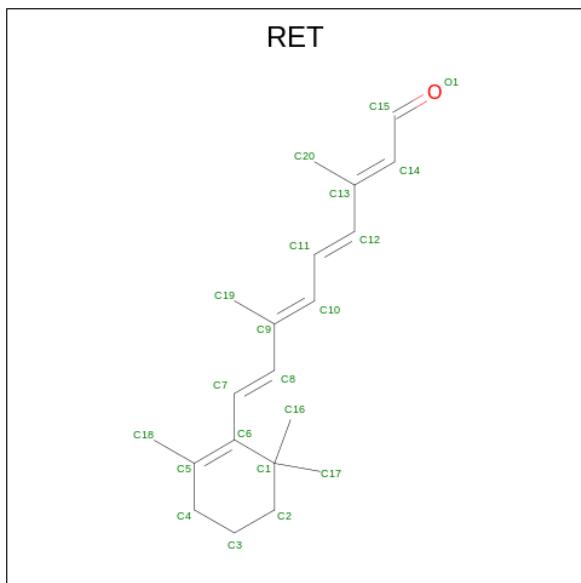
- Molecule 5 is a protein called Medium-wave-sensitive opsin 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	R	298	Total	C	N	O	S	0	0
			2327	1552	373	382	20		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	129	GLN	GLU	engineered mutation	UNP P04001

- Molecule 6 is RETINAL (CCD ID: RET) (formula: C<sub>20</sub>H<sub>28</sub>O) (labeled as "Ligand of Interest" by depositor).

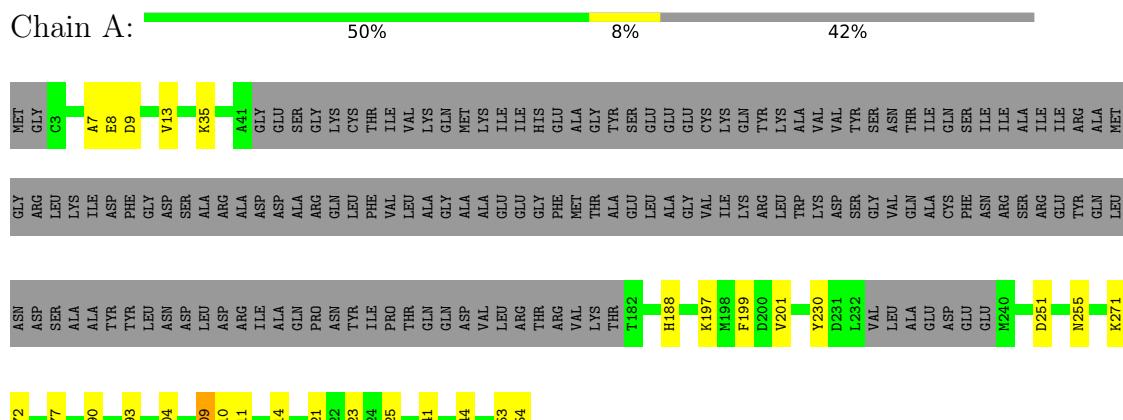


Mol	Chain	Residues	Atoms	AltConf
6	R	1	Total 20 20	0

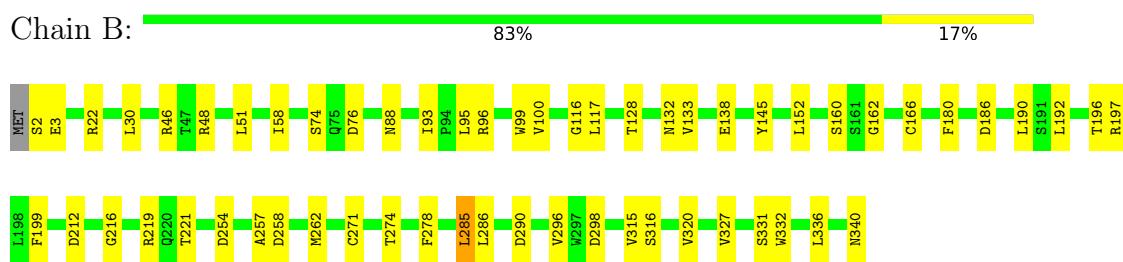
### 3 Residue-property plots

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.

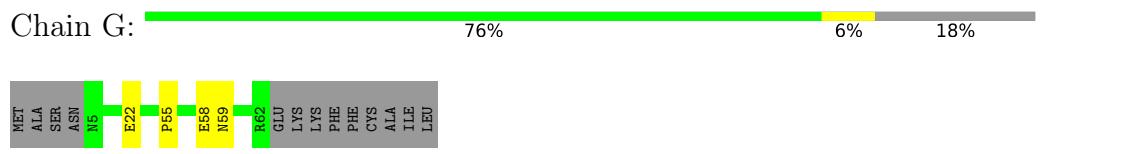
- Molecule 1: Guanine nucleotide-binding protein G(i) subunit alpha-1



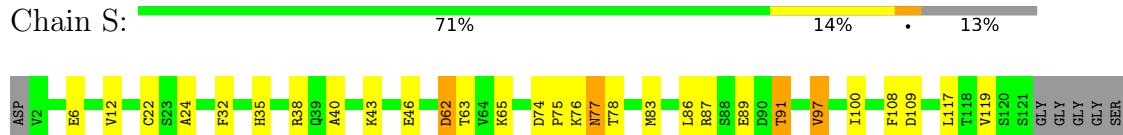
- Molecule 2: Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1

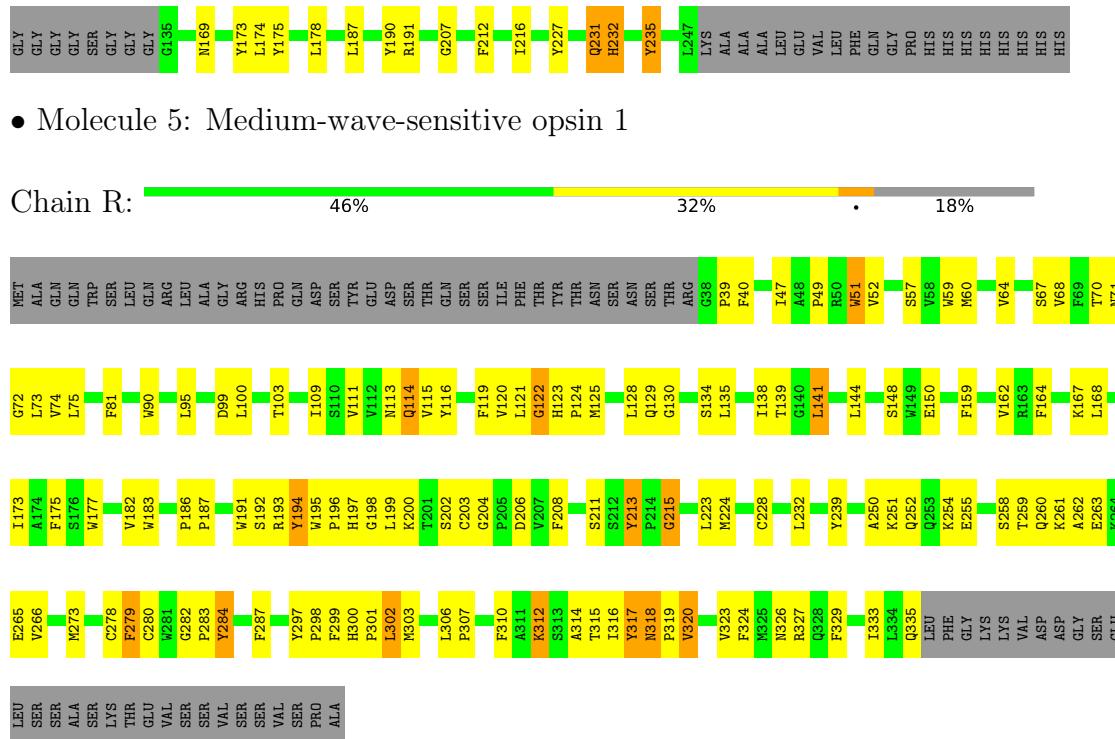


- Molecule 3: Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2



- Molecule 4: scFv Recombinant Human Monoclonal Antibody (scFv16)





## 4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	577973	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TECNAI 10	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	54	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor

## 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: RET

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	0.56	0/1599	0.76	1/2161 (0.0%)
2	B	0.58	0/2646	0.72	1/3587 (0.0%)
3	G	0.32	0/444	0.42	0/601
4	S	0.62	0/1795	1.09	6/2438 (0.2%)
5	R	0.93	1/2407 (0.0%)	1.17	12/3301 (0.4%)
All	All	0.69	1/8891 (0.0%)	0.94	20/12088 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	R	312	LYS	C-N	-5.48	1.26	1.33

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	S	231	GLN	N-CA-C	31.70	145.91	111.36
5	R	194	TYR	N-CA-C	-10.48	97.01	110.53
4	S	191	ARG	N-CA-C	-7.64	98.07	108.38
5	R	320	VAL	N-CA-C	7.52	117.64	110.42
4	S	231	GLN	CB-CA-C	-7.17	98.66	110.85

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	1570	0	1498	20	0
2	B	2599	0	2502	35	0
3	G	438	0	443	3	0
4	S	1751	0	1673	26	0
5	R	2327	0	2309	135	0
6	R	20	0	27	9	0
All	All	8705	0	8452	209	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 12.

The worst 5 of 209 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)
5:R:51:TRP:CD1	5:R:51:TRP:C	2.36	1.02
5:R:280:CYS:SG	5:R:314:ALA:HB2	2.06	0.95
5:R:193:ARG:CB	5:R:213:TYR:CE2	2.60	0.84
5:R:51:TRP:C	5:R:51:TRP:HD1	1.82	0.83
5:R:99:ASP:HB3	5:R:315:THR:HG23	1.62	0.82

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 PDB entries followed by that with respect to all EM entries.

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	199/354 (56%)	190 (96%)	9 (4%)	0	100 100
2	B	337/340 (99%)	315 (94%)	22 (6%)	0	100 100
3	G	56/71 (79%)	56 (100%)	0	0	100 100
4	S	229/267 (86%)	220 (96%)	7 (3%)	2 (1%)	14 26

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
5	R	296/364 (81%)	286 (97%)	10 (3%)	0	100 100
All	All	1117/1396 (80%)	1067 (96%)	48 (4%)	2 (0%)	45 62

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	S	62	ASP
4	S	77	ASN

### 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 PDB entries followed by that with respect to all EM entries.

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	159/306 (52%)	156 (98%)	3 (2%)	52 74
2	B	280/283 (99%)	279 (100%)	1 (0%)	89 95
3	G	46/58 (79%)	46 (100%)	0	100 100
4	S	185/216 (86%)	181 (98%)	4 (2%)	47 70
5	R	244/310 (79%)	228 (93%)	16 (7%)	14 26
All	All	914/1173 (78%)	890 (97%)	24 (3%)	42 65

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	R	144	LEU
5	R	223	LEU
5	R	213	TYR
5	R	232	LEU
4	S	232	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
3	G	24	ASN
5	R	114	GLN
5	R	331	ASN
5	R	129	GLN
2	B	91	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 oligosaccharides in this entry.

### 5.6 Ligand geometry [\(i\)](#)

1 ligand is 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	$\# Z  > 2$
6	RET	R	401	5	20,20,21	2.93	5 (25%)	27,27,28	1.08	2 (7%)

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
6	RET	R	401	5	-	2/13/30/31	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	R	401	RET	C14-C13	11.38	1.42	1.33
6	R	401	RET	C10-C9	3.89	1.40	1.35
6	R	401	RET	C15-C14	-2.64	1.39	1.49
6	R	401	RET	C5-C6	2.39	1.38	1.34
6	R	401	RET	C11-C12	2.19	1.40	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	R	401	RET	C11-C10-C9	-2.73	123.41	127.31
6	R	401	RET	C7-C8-C9	-2.09	123.07	126.23

There are no chirality outliers.

All (2) torsion outliers are listed below:

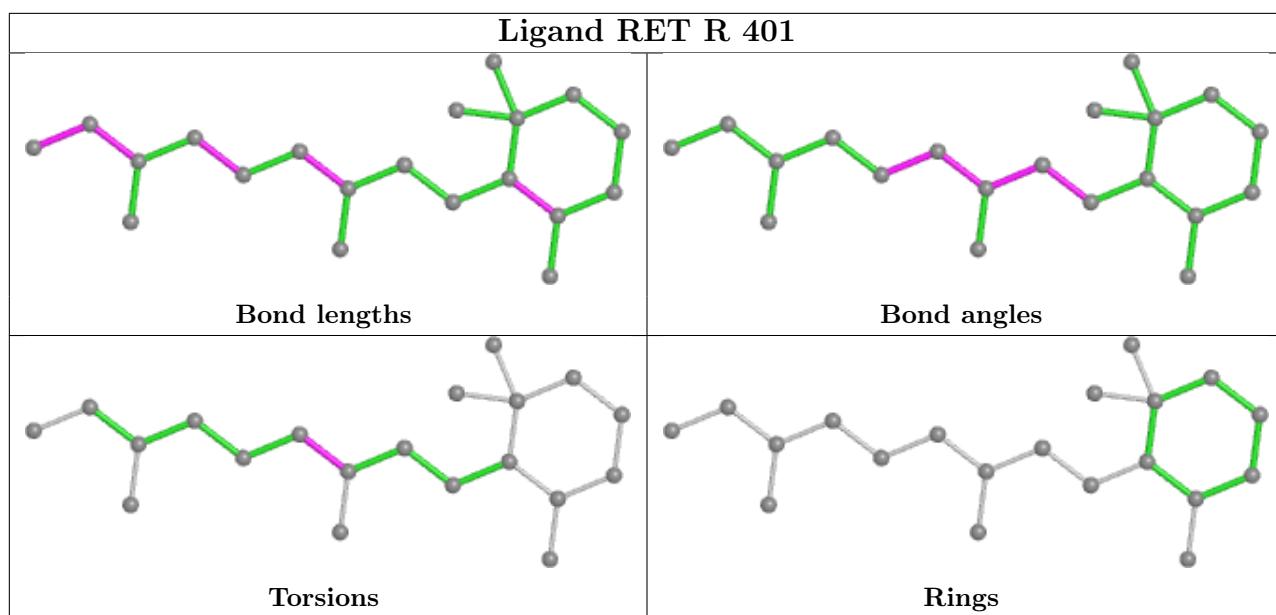
Mol	Chain	Res	Type	Atoms
6	R	401	RET	C11-C10-C9-C8
6	R	401	RET	C11-C10-C9-C19

There are no ring outliers.

1 monomer is involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	R	401	RET	9	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 than 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.



## 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.