



wwPDB EM Validation Summary Report i

Jul 16, 2025 – 10:55 AM JST

PDB ID : 8Y02 / pdb_00008y02
EMDB ID : EMD-38801
Title : Cryo-EM structure of Short-wave-sensitive opsin 1
Authors : Peng, Q.; Jiang, H.H.; Cheng, X.Y.; Li, J.; Zhang, J.
Deposited on : 2024-01-21
Resolution : 2.61 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
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:

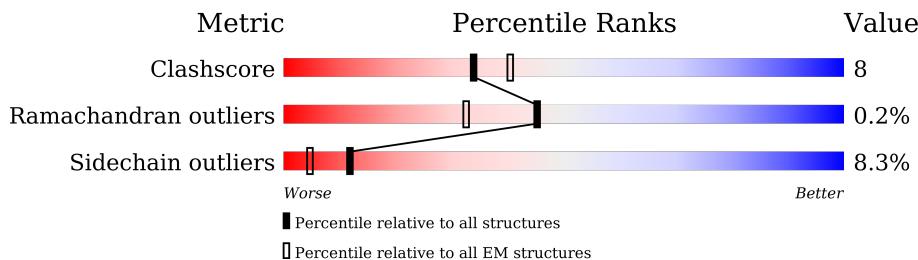
EMDB validation analysis	: FAILED
Mogul	: 1.8.5 (274361), CSD as541be (2020)
MolProbity	: 4-5-2 with Phenix2.0rc1
buster-report	: 1.1.7 (2018)
Percentile statistics	: 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	: FAILED
Ideal geometry (proteins)	: Engh & Huber (2001)
Ideal 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 2.61 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 ≥ 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 $\leq 5\%$



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

There are 6 unique types of molecules in this entry. The entry contains 8156 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 Short-wave-sensitive opsin 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	R	258	2002	1339	316	327	20	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	1	MET	-	initiating methionine	UNP P03999
R	2	ARG	-	expression tag	UNP P03999
R	3	LYS	-	expression tag	UNP P03999
R	110	GLN	GLU	engineered mutation	UNP P03999

- Molecule 2 is a protein called Guanine nucleotide-binding protein G(i) subunit alpha-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	A	208	1582	1014	267	289	12	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	B	325	2469	1526	443	479	21	0	0

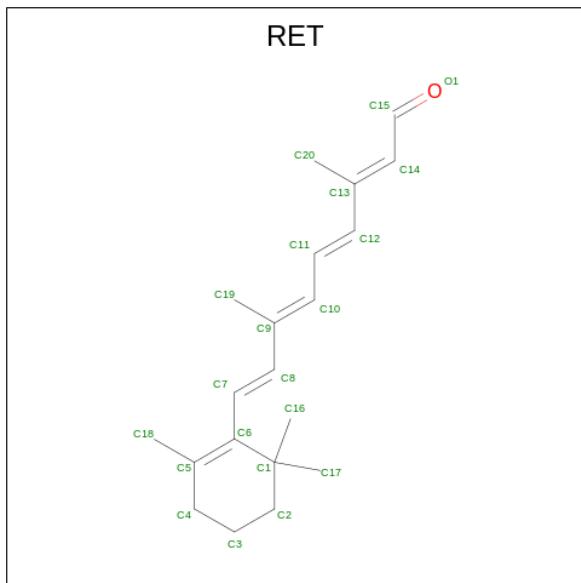
- Molecule 4 is a protein called Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	G	48	Total	C	N	O	S	0
			356	225	60	69	2	0

- Molecule 5 is a protein called scFv Recombinant Human Monoclonal Antibody (scFv16).

Mol	Chain	Residues	Atoms				AltConf	Trace
5	S	230	Total	C	N	O	S	0
			1727	1101	289	327	10	0

- Molecule 6 is RETINAL (CCD ID: RET) (formula: C₂₀H₂₈O) (labeled as "Ligand of Interest" by depositor).

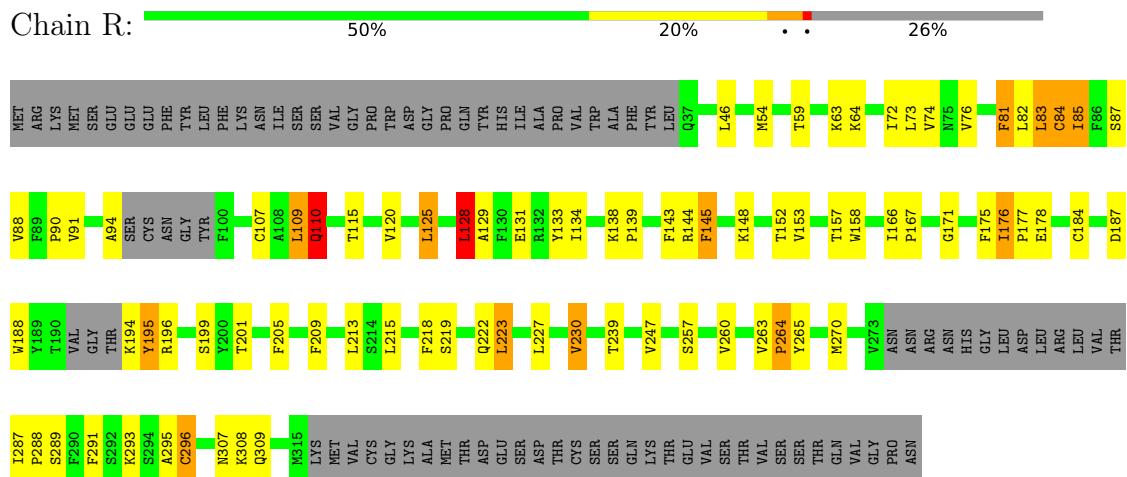


Mol	Chain	Residues	Atoms		AltConf
6	R	1	Total	C	0
			20	20	

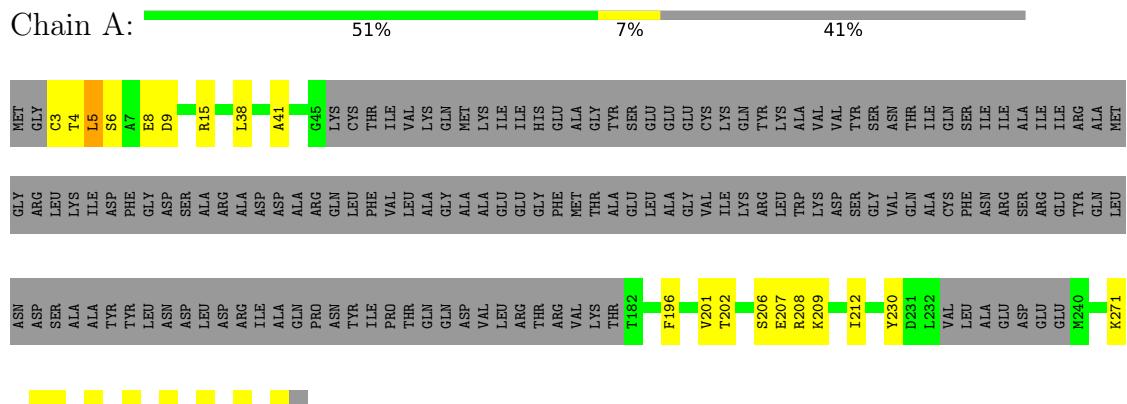
3 Residue-property plots

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: Short-wave-sensitive opsin 1



- Molecule 2: Guanine nucleotide-binding protein G(i) subunit alpha-1



- Molecule 3: Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1

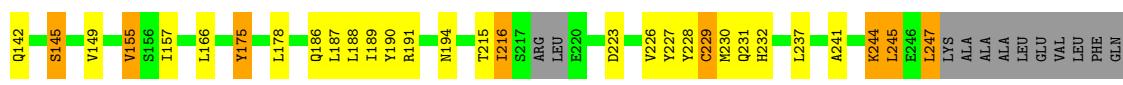




- Molecule 4: Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2



- Molecule 5: 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	405968	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 ($e^-/\text{\AA}^2$)	51.9	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	R	0.98	1/2060 (0.0%)	1.41	11/2797 (0.4%)
2	A	0.73	1/1610 (0.1%)	0.84	2/2174 (0.1%)
3	B	0.92	0/2516	0.97	4/3416 (0.1%)
4	G	0.71	0/362	0.90	0/492
5	S	0.97	5/1770 (0.3%)	1.11	2/2403 (0.1%)
All	All	0.91	7/8318 (0.1%)	1.10	19/11282 (0.2%)

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
3	B	0	1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	S	230	MET	C-O	-8.90	1.13	1.23
5	S	228	TYR	C-O	-8.44	1.13	1.23
5	S	47	TRP	C-O	-7.72	1.14	1.23
5	S	229	CYS	C-O	-6.43	1.15	1.23
1	R	295	ALA	C-O	-5.41	1.17	1.24

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	R	81	PHE	N-CA-C	-9.74	100.94	113.12
1	R	138	LYS	CB-CA-C	9.53	119.67	110.71
1	R	167	PRO	N-CA-C	7.53	119.89	110.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	R	128	LEU	N-CA-C	-7.14	102.69	111.40
1	R	188	TRP	N-CA-C	-7.04	105.31	114.04

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	219	ARG	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	R	2002	0	1986	40	0
2	A	1582	0	1513	13	0
3	B	2469	0	2363	42	0
4	G	356	0	347	6	0
5	S	1727	0	1644	33	0
6	R	20	0	27	5	0
All	All	8156	0	7880	129	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 8.

The worst 5 of 129 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:26:ALA:HB2	3:B:259:GLN:HE22	1.52	0.73
1:R:88:VAL:HG22	1:R:293:LYS:HD2	1.72	0.72
4:G:32:LYS:O	4:G:32:LYS:HD3	1.90	0.71
3:B:271:CYS:HB2	3:B:290:ASP:HB2	1.73	0.71
3:B:126:LEU:HA	3:B:133:VAL:HG12	1.72	0.69

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	R	250/348 (72%)	232 (93%)	16 (6%)	2 (1%)	16 32
2	A	202/354 (57%)	196 (97%)	6 (3%)	0	100 100
3	B	323/340 (95%)	303 (94%)	20 (6%)	0	100 100
4	G	46/71 (65%)	44 (96%)	2 (4%)	0	100 100
5	S	224/267 (84%)	214 (96%)	10 (4%)	0	100 100
All	All	1045/1380 (76%)	989 (95%)	54 (5%)	2 (0%)	45 65

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	R	143	PHE
1	R	171	GLY

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	R	209/300 (70%)	176 (84%)	33 (16%)	2 3
2	A	160/306 (52%)	154 (96%)	6 (4%)	28 52
3	B	264/283 (93%)	248 (94%)	16 (6%)	15 32
4	G	36/58 (62%)	32 (89%)	4 (11%)	5 9
5	S	183/216 (85%)	171 (93%)	12 (7%)	14 28
All	All	852/1163 (73%)	781 (92%)	71 (8%)	12 18

5 of 71 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	G	32	LYS
5	S	91	THR
5	S	186	GLN
1	R	230	VAL
1	R	227	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	119	ASN
3	B	340	ASN
5	S	194	ASN
4	G	59	ASN
3	B	259	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 [\(i\)](#)

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	1	20,20,21	2.59	5 (25%)	27,27,28	0.84	1 (3%)

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	1	-	0/13/30/31	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	9.72	1.41	1.33
6	R	401	RET	C10-C9	3.58	1.40	1.35
6	R	401	RET	C5-C6	2.55	1.38	1.34
6	R	401	RET	C15-C14	-2.52	1.40	1.49
6	R	401	RET	C11-C12	2.03	1.39	1.34

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	R	401	RET	C20-C13-C12	2.64	122.23	118.08

There are no chirality outliers.

There are no torsion outliers.

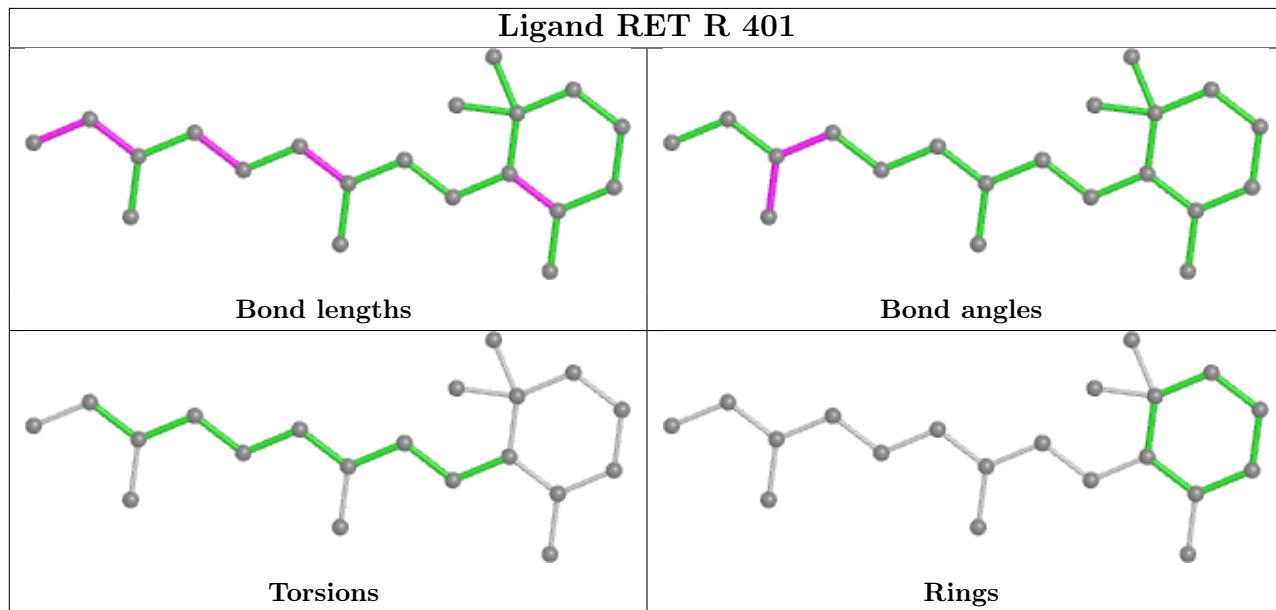
There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	R	401	RET	5	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.