



# Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 09:34 am BST

PDB ID : 6C7P  
Title : Crystal structure of D477G ACO/RPE65 chimera, monoclinic crystal form  
Authors : Kiser, P.D.; Shi, W.  
Deposited on : 2018-01-23  
Resolution : 2.60 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

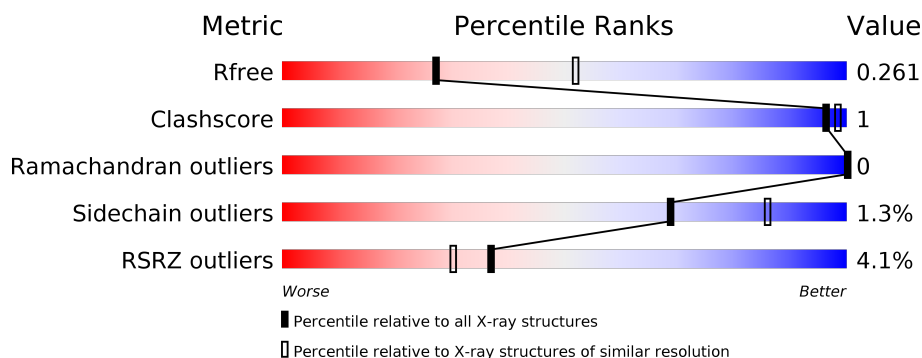
# 1 Overall quality at a glance

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))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (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 on the lower 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 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	490	<div> <div>4%</div> <div> <div></div> <div>94%</div> <div></div> </div> <div></div> </div>
1	B	490	<div> <div>3%</div> <div> <div></div> <div>93%</div> <div></div> </div> <div></div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7546 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 Apocarotenoid-15,15'-oxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	474	Total	C	N	O	S	0	0	0
			3738	2404	644	680	10			
1	A	476	Total	C	N	O	S	0	0	0
			3753	2412	647	684	10			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	44	LEU	PRO	engineered mutation	UNP P74334
B	45	TRP	ASP	engineered mutation	UNP P74334
B	431	SER	PRO	engineered mutation	UNP P74334
B	432	HIS	ARG	engineered mutation	UNP P74334
B	435	ALA	GLY	engineered mutation	UNP P74334
B	436	LEU	VAL	engineered mutation	UNP P74334
B	437	GLU	ALA	engineered mutation	UNP P74334
B	442	VAL	TRP	engineered mutation	UNP P74334
B	461	ASN	ASP	engineered mutation	UNP P74334
B	463	LYS	GLN	engineered mutation	UNP P74334
A	44	LEU	PRO	engineered mutation	UNP P74334
A	45	TRP	ASP	engineered mutation	UNP P74334
A	431	SER	PRO	engineered mutation	UNP P74334
A	432	HIS	ARG	engineered mutation	UNP P74334
A	435	ALA	GLY	engineered mutation	UNP P74334
A	436	LEU	VAL	engineered mutation	UNP P74334
A	437	GLU	ALA	engineered mutation	UNP P74334
A	442	VAL	TRP	engineered mutation	UNP P74334
A	461	ASN	ASP	engineered mutation	UNP P74334
A	463	LYS	GLN	engineered mutation	UNP P74334

- Molecule 2 is FE (II) ION (three-letter code: FE2) (formula: Fe).

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

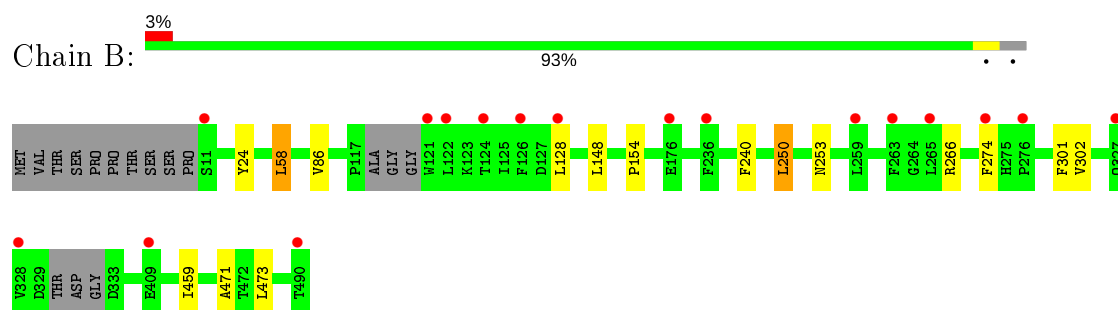
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	27	Total 27	O 27	0	0
3	A	26	Total 26	O 26	0	0

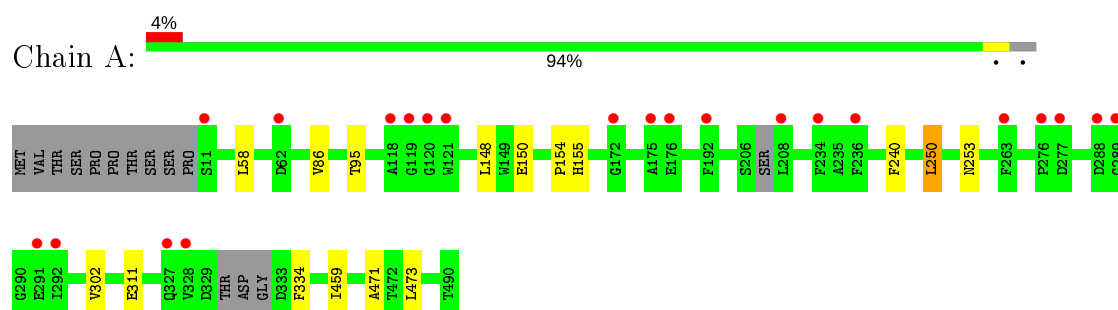
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Apocarotenoid-15,15'-oxygenase



- Molecule 1: Apocarotenoid-15,15'-oxygenase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	191.43 Å   56.17 Å   136.91 Å 90.00°   133.82°   90.00°	Depositor
Resolution (Å)	41.76 – 2.60 41.76 – 2.60	Depositor EDS
% Data completeness (in resolution range)	97.3 (41.76-2.60) 97.3 (41.76-2.60)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.00 (at 2.61 Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.220 , 0.260 0.227 , 0.261	Depositor DCC
$R_{free}$ test set	1588 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	81.4	Xtriage
Anisotropy	0.323	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 42.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.000 for h+2*k,-h-l 0.011 for h,-k,-h-l 0.009 for -h-2*k,-k,l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7546	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.88% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: FE2

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	0.37	0/3863	0.56	0/5255
1	B	0.37	0/3848	0.56	0/5237
All	All	0.37	0/7711	0.56	0/10492

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 [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	3753	0	3656	7	0
1	B	3738	0	3634	8	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	26	0	0	0	0
3	B	27	0	0	0	0
All	All	7546	0	7290	15	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.

All (15) 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:459:ILE:HD12	1:B:471:ALA:HB3	1.95	0.49
1:A:459:ILE:HD12	1:A:471:ALA:HB3	1.95	0.49
1:B:86:VAL:HG11	1:B:473:LEU:CD2	2.44	0.48
1:A:86:VAL:HG11	1:A:473:LEU:CD2	2.44	0.48
1:B:250:LEU:HD13	1:B:302:VAL:HG11	1.95	0.48
1:B:148:LEU:HD23	1:B:154:PRO:HB3	1.96	0.47
1:A:148:LEU:HD23	1:A:154:PRO:HB3	1.97	0.46
1:A:250:LEU:HD13	1:A:302:VAL:HG11	1.97	0.46
1:A:95:THR:HG22	1:A:155:HIS:CE1	2.51	0.46
1:B:86:VAL:HG11	1:B:473:LEU:HD23	1.99	0.45
1:A:86:VAL:HG11	1:A:473:LEU:HD23	2.01	0.43
1:B:24:TYR:HB3	1:B:58:LEU:HD22	2.01	0.43
1:B:128:LEU:N	1:B:128:LEU:HD12	2.35	0.41
1:A:58:LEU:HD21	1:A:334:PHE:CZ	2.56	0.41
1:B:274:PHE:CE2	1:B:301:PHE:HB2	2.56	0.41

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	470/490 (96%)	456 (97%)	14 (3%)	0	100	100
1	B	468/490 (96%)	453 (97%)	15 (3%)	0	100	100
All	All	938/980 (96%)	909 (97%)	29 (3%)	0	100	100

There are no Ramachandran outliers to report.



### 5.3.2 Protein sidechains ⓘ

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	398/411 (97%)	393 (99%)	5 (1%)	69	86
1	B	396/411 (96%)	391 (99%)	5 (1%)	69	86
All	All	794/822 (97%)	784 (99%)	10 (1%)	69	86

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	58	LEU
1	B	240	PHE
1	B	250	LEU
1	B	253	ASN
1	B	266	ARG
1	A	150	GLU
1	A	240	PHE
1	A	250	LEU
1	A	253	ASN
1	A	311	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	153	GLN
1	B	253	ASN
1	B	397	ASN
1	A	153	GLN
1	A	253	ASN
1	A	397	ASN

### 5.3.3 RNA ⓘ

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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 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.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	476/490 (97%)	0.14	22 (4%) 32 26	50, 77, 108, 136	0
1	B	474/490 (96%)	0.10	17 (3%) 42 35	53, 77, 106, 133	0
All	All	950/980 (96%)	0.12	39 (4%) 37 30	50, 77, 107, 136	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	328	VAL	6.2
1	A	236	PHE	5.7
1	B	327	GLN	4.6
1	A	208	LEU	4.6
1	B	236	PHE	4.5
1	B	276	PRO	4.0
1	A	234	PHE	3.9
1	B	128	LEU	3.7
1	B	259	LEU	3.5
1	A	118	ALA	3.2
1	A	119	GLY	3.2
1	A	288	ASP	3.1
1	A	328	VAL	3.1
1	B	126	PHE	3.1
1	B	490	THR	3.1
1	A	11	SER	3.1
1	B	11	SER	3.0
1	A	120	GLY	2.9
1	B	121	TRP	2.9
1	A	176	GLU	2.8
1	A	292	ILE	2.8
1	B	274	PHE	2.8
1	A	291	GLU	2.7
1	A	327	GLN	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	289	GLY	2.6
1	B	263	PHE	2.5
1	A	175	ALA	2.5
1	A	172	GLY	2.4
1	B	409	GLU	2.4
1	A	121	TRP	2.4
1	A	263	PHE	2.4
1	B	265	LEU	2.3
1	B	122	LEU	2.3
1	A	276	PRO	2.2
1	B	176	GLU	2.1
1	A	277	ASP	2.1
1	B	124	THR	2.1
1	A	62	ASP	2.1
1	A	192	PHE	2.1

## 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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	FE2	A	501	1/1	0.90	0.18	81,81,81,81	0
2	FE2	B	501	1/1	0.99	0.17	77,77,77,77	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.