



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 21, 2017 – 01:39 AM EDT

PDB ID : 5MSW  
Title : Structure of the A-PCP didomain of carboxylic acid reductase (CAR) from  
Segniliparus rugosus in complex with AMP  
Authors : Gahloth, D.; Leys, D.  
Deposited on : unknown  
Resolution : 2.33 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

<http://wwpdb.org/validation/2016/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  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20029824  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029824

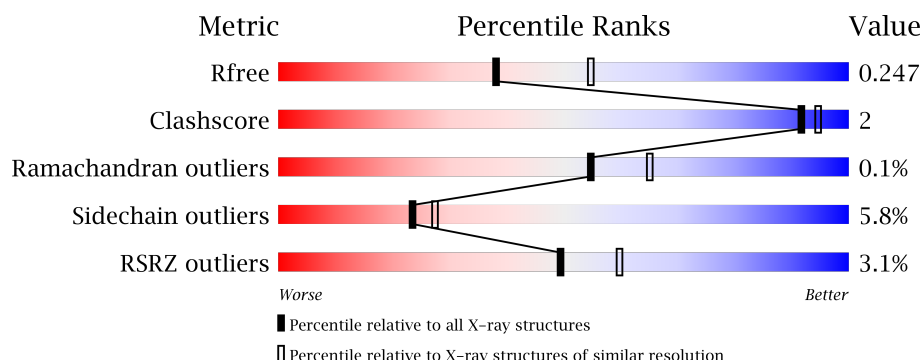
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.33 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))
$R_{free}$	100719	1570 (2.36-2.32)
Clashscore	112137	1673 (2.36-2.32)
Ramachandran outliers	110173	1654 (2.36-2.32)
Sidechain outliers	110143	1655 (2.36-2.32)
RSRZ outliers	101464	1576 (2.36-2.32)

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. 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	1188	<div> <div>2%</div> <div>54%</div> <div>6%</div> <div>40%</div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5510 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 Thioester reductase domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	715	5431	3439	929	1056	7	0	0	0

- Molecule 2 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula:  $C_{10}H_{14}N_5O_7P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	23	10	5	7	1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	56	Total 56 O 56	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	131.42Å 65.33Å 101.71Å 90.00° 110.80° 90.00°	Depositor
Resolution (Å)	95.08 – 2.33 34.70 – 2.33	Depositor EDS
% Data completeness (in resolution range)	99.0 (95.08-2.33) 99.1 (34.70-2.33)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.194 , 0.243 0.202 , 0.247	Depositor DCC
$R_{free}$ test set	1738 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	42.5	Xtriage
Anisotropy	0.084	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 28.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5510	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.73% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: AMP

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	0.74	0/5535	0.89	11/7527 (0.1%)

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	105	ARG	NE-CZ-NH2	-9.80	115.40	120.30
1	A	733	ARG	NE-CZ-NH1	7.34	123.97	120.30
1	A	45	ARG	NE-CZ-NH1	6.72	123.66	120.30
1	A	441	ARG	NE-CZ-NH1	6.64	123.62	120.30
1	A	136	ASP	CB-CG-OD1	6.29	123.96	118.30
1	A	659	ARG	NE-CZ-NH1	6.18	123.39	120.30
1	A	621	ASN	CB-CA-C	-5.69	99.02	110.40
1	A	393	ARG	NE-CZ-NH1	5.32	122.96	120.30
1	A	73	ARG	NE-CZ-NH2	-5.18	117.71	120.30
1	A	105	ARG	NE-CZ-NH1	5.18	122.89	120.30
1	A	39	ASP	CB-CG-OD1	5.01	122.81	118.30

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	5431	0	5412	21	0
2	A	23	0	12	0	0
3	A	56	0	0	0	0
All	All	5510	0	5424	21	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 2.

All (21) 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:A:112:GLU:OE2	1:A:239:ARG:NH1	2.34	0.59
1:A:159:THR:OG1	1:A:162:THR:HG23	2.06	0.56
1:A:101:GLU:O	1:A:105:ARG:HG2	2.09	0.53
1:A:455:LEU:N	1:A:455:LEU:HD12	2.27	0.49
1:A:711:PHE:CE2	1:A:715:ILE:HD11	2.48	0.49
1:A:100:ARG:O	1:A:104:GLU:HG2	2.13	0.47
1:A:735:VAL:O	1:A:739:ILE:HG23	2.14	0.47
1:A:292:ASN:O	1:A:441:ARG:HD3	2.16	0.46
1:A:625:SER:HB3	1:A:631:LEU:HD21	1.97	0.45
1:A:741:LYS:C	1:A:742:GLU:HG2	2.36	0.45
1:A:392:LEU:HD23	1:A:396:LEU:HD12	1.99	0.45
1:A:632:ARG:HB3	1:A:633:PRO:HD3	2.00	0.43
1:A:105:ARG:NH2	1:A:136:ASP:OD1	2.40	0.43
1:A:409:ALA:HB1	1:A:410:PRO:HD2	2.00	0.43
1:A:446:VAL:HG12	1:A:448:PRO:O	2.19	0.43
1:A:80:ASP:HB3	1:A:83:THR:HG22	2.01	0.42
1:A:411:MET:HG2	1:A:416:THR:OG1	2.19	0.42
1:A:507:ASP:OD1	1:A:522:ARG:HG2	2.21	0.41
1:A:617:PHE:HB3	1:A:623:ILE:HB	2.02	0.41
1:A:376:ILE:HD12	1:A:388:ILE:HD11	2.03	0.40
1:A:313:LEU:HD21	1:A:322:LEU:HD23	2.03	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	707/1188 (60%)	690 (98%)	16 (2%)	1 (0%)	55	65

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	667	GLU

### 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	568/947 (60%)	535 (94%)	33 (6%)	23	27

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	83	THR
1	A	87	VAL
1	A	137	TYR
1	A	151	SER
1	A	152	VAL
1	A	162	THR
1	A	163	LEU
1	A	169	GLU
1	A	198	LEU
1	A	205	GLN
1	A	220	ARG
1	A	256	ASP
1	A	265	THR
1	A	266	SER
1	A	273	LYS
1	A	311	MET
1	A	323	MET
1	A	335	ILE

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	337	LYS
1	A	338	SER
1	A	340	LEU
1	A	360	ARG
1	A	394	GLU
1	A	412	SER
1	A	491	GLU
1	A	527	LEU
1	A	586	SER
1	A	688	GLU
1	A	715	ILE
1	A	731	ASN
1	A	733	ARG
1	A	739	ILE
1	A	742	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

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	$\# Z  > 2$
2	AMP	A	1201	-	22,25,25	1.14	3 (13%)	24,38,38	1.90	4 (16%)

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
2	AMP	A	1201	-	-	0/6/26/26	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1201	AMP	O4'-C1'	2.11	1.44	1.41
2	A	1201	AMP	C2-N3	2.22	1.35	1.32
2	A	1201	AMP	C5-C4	3.18	1.47	1.40

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1201	AMP	N3-C2-N1	-7.27	122.53	128.86
2	A	1201	AMP	C4-C5-N7	-2.77	106.74	109.41
2	A	1201	AMP	O3P-P-O5'	-2.05	101.29	106.73
2	A	1201	AMP	C2-N1-C6	2.18	122.58	118.77

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 ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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, 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	715/1188 (60%)	0.10	22 (3%) 49 60	23, 43, 84, 108	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	735	VAL	4.2
1	A	692	GLU	3.8
1	A	723	SER	3.0
1	A	739	ILE	3.0
1	A	317	ALA	3.0
1	A	741	LYS	2.8
1	A	732	LEU	2.7
1	A	318	GLY	2.6
1	A	29	ILE	2.6
1	A	734	SER	2.5
1	A	697	ASP	2.5
1	A	294	PHE	2.4
1	A	716	PHE	2.3
1	A	717	GLN	2.3
1	A	728	ALA	2.2
1	A	376	ILE	2.2
1	A	736	ALA	2.1
1	A	738	HIS	2.1
1	A	320	ILE	2.1
1	A	696	SER	2.1
1	A	722	VAL	2.0
1	A	309	ASN	2.0

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	AMP	A	1201	23/23	0.97	0.14	-0.71	29,39,55,59	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.