



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 12, 2024 – 06:38 PM EST

PDB ID : 3SJR
Title : Crystal structure of conserved unknown function protein CV_1783 from *Chromobacterium violaceum* ATCC 12472
Authors : Chang, C.; Hatzos-Skintges, C.; Clancy, S.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on : 2011-06-21
Resolution : 2.94 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

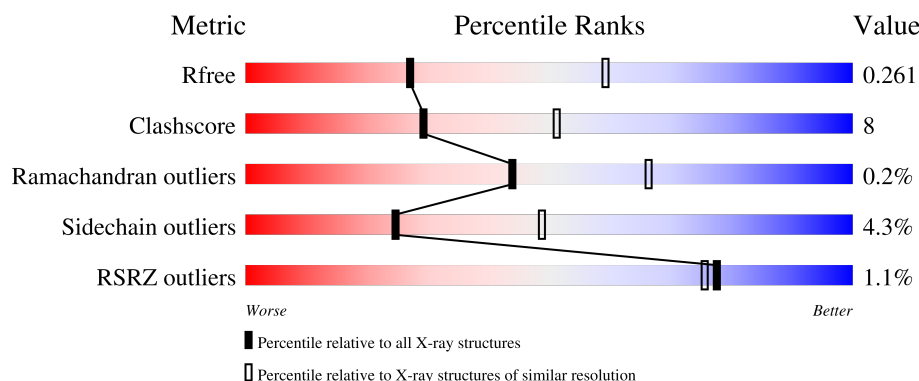
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.94 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}	164625	1067 (2.96-2.92)
Clashscore	180529	1122 (2.96-2.92)
Ramachandran outliers	177936	1075 (2.96-2.92)
Sidechain outliers	177891	1075 (2.96-2.92)
RSRZ outliers	164620	1067 (2.96-2.92)

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 $\leq 5\%$. 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	175	 2% 60% 11% 28%
1	B	175	 1% 58% 7% 33%
1	C	175	 43% 22% 33%
1	D	175	 50% 13% 37%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3652 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 Uncharacterized protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	126	Total	C	N	O	S	Se	0	0	0
			974	615	175	180	1	3			
1	B	117	Total	C	N	O	S	Se	0	0	0
			904	570	163	167	1	3			
1	C	117	Total	C	N	O	S	Se	0	0	0
			909	572	163	170	1	3			
1	D	110	Total	C	N	O	S	Se	0	0	0
			861	543	156	158	1	3			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q7NX46
A	-1	ASN	-	expression tag	UNP Q7NX46
A	0	ALA	-	expression tag	UNP Q7NX46
B	-2	SER	-	expression tag	UNP Q7NX46
B	-1	ASN	-	expression tag	UNP Q7NX46
B	0	ALA	-	expression tag	UNP Q7NX46
C	-2	SER	-	expression tag	UNP Q7NX46
C	-1	ASN	-	expression tag	UNP Q7NX46
C	0	ALA	-	expression tag	UNP Q7NX46
D	-2	SER	-	expression tag	UNP Q7NX46
D	-1	ASN	-	expression tag	UNP Q7NX46
D	0	ALA	-	expression tag	UNP Q7NX46

- Molecule 2 is water.

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

Continued on next page...

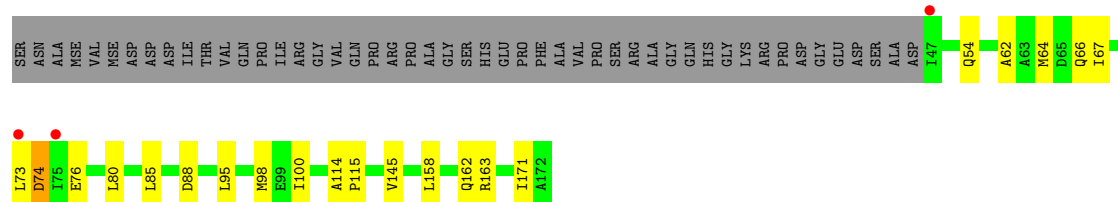
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	1	Total	O	0	0
			1	1		
2	D	1	Total	O	0	0
			1	1		

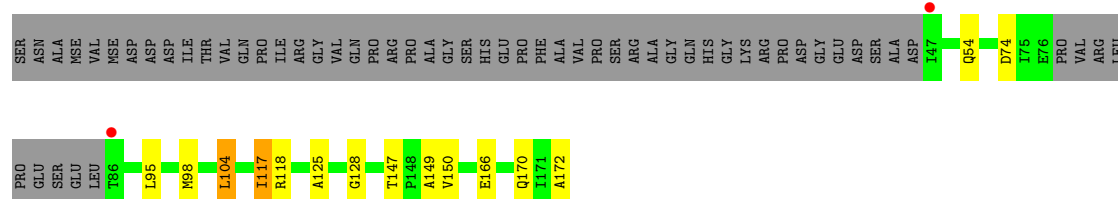
3 Residue-property plots [i](#)

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 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.

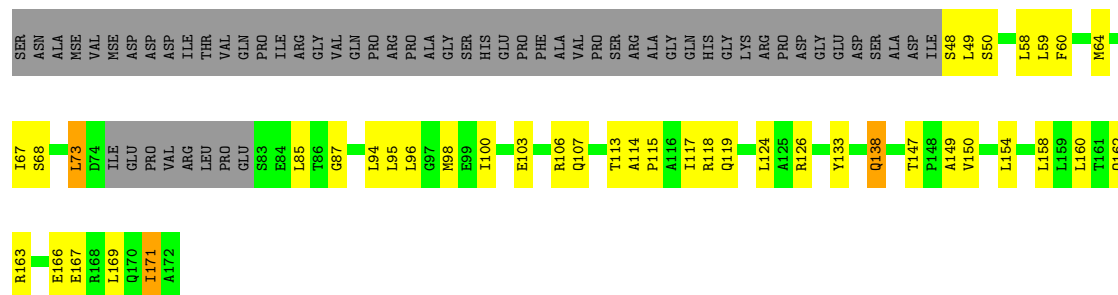
- Molecule 1: Uncharacterized protein



- Molecule 1: Uncharacterized protein



- Molecule 1: Uncharacterized protein



- Molecule 1: Uncharacterized protein



SER	ASN	ALA	NSE	VAL	NSE	ASP	ASP	ASP	ILE	THR	VAL	GLN	PRO	PRO	ILE	ARG	GLY	VAL	GLN	PRO	ARG	PRO	ALA	GLY	SER	HIS	GLU	PRO	PHE	ALA	VAL	VAL	PRO	SER	ARG	ALA	GLY	GLN	HIS	GLY	LYS	ARG	PRO	ASP	GLY	GLU	GLY	ASP	SER	SER	ALA	ASP	ILE	SER	L49	S50	Q51	Q54	L59	S68
R69	L70	L73	D74	ILE	GLU	PRO	VAL	ARG	ANG	LEU	PRO	GLU	SER	GLU	LEU	THR	GLY	D88	L95	M98	D111	P115	A116	I117	R118	Q119	A120	F121	A122	P123	G128	R131	T147	L154	L157	F165	E166	E167	I171	ALA																				

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	95.37Å 109.47Å 151.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.94 50.00 – 2.94	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.00-2.94) 99.3 (50.00-2.94)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.92 (at 2.95Å)	Xtriage
Refinement program	REFMAC 5.5.0109, CNS	Depositor
R, R_{free}	0.226 , 0.267 0.224 , 0.261	Depositor DCC
R_{free} test set	867 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	81.6	Xtriage
Anisotropy	0.348	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 41.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3652	wwPDB-VP
Average B, all atoms (Å ²)	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.02% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

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	0.51	0/987	0.62	0/1334
1	B	0.50	0/914	0.62	0/1232
1	C	0.48	0/919	0.61	0/1238
1	D	0.46	0/871	0.57	0/1173
All	All	0.49	0/3691	0.61	0/4977

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	974	0	975	14	0
1	B	904	0	904	11	0
1	C	909	0	911	29	0
1	D	861	0	860	16	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
All	All	3652	0	3650	62	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 62 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:C:67:ILE:HG12	1:D:128:GLY:HA3	1.78	0.66
1:D:116:ALA:HA	1:D:119:GLN:HG3	1.81	0.62
1:D:115:PRO:O	1:D:119:GLN:HG2	1.99	0.61
1:C:126:ARG:HH21	1:C:166:GLU:HG3	1.65	0.61
1:D:122:ALA:HB3	1:D:123:PRO:HD3	1.84	0.59

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	124/175 (71%)	120 (97%)	3 (2%)	1 (1%)	16	38
1	B	113/175 (65%)	108 (96%)	5 (4%)	0	100	100
1	C	113/175 (65%)	111 (98%)	2 (2%)	0	100	100
1	D	106/175 (61%)	100 (94%)	6 (6%)	0	100	100
All	All	456/700 (65%)	439 (96%)	16 (4%)	1 (0%)	44	66

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	74	ASP

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	99/136 (73%)	95 (96%)	4 (4%)	27	51
1	B	91/136 (67%)	88 (97%)	3 (3%)	33	57
1	C	93/136 (68%)	88 (95%)	5 (5%)	18	40
1	D	87/136 (64%)	83 (95%)	4 (5%)	23	46
All	All	370/544 (68%)	354 (96%)	16 (4%)	25	49

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	154	LEU
1	D	147	THR
1	C	73	LEU
1	D	111	ASP
1	C	49	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	142	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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 [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, 95th 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 ²)	Q<0.9
1	A	123/175 (70%)	-0.20	3 (2%) 59 57	62, 85, 112, 132	0
1	B	114/175 (65%)	-0.22	2 (1%) 67 65	63, 85, 118, 142	0
1	C	114/175 (65%)	-0.16	0 100 100	66, 88, 120, 142	0
1	D	107/175 (61%)	-0.25	0 100 100	66, 93, 127, 143	0
All	All	458/700 (65%)	-0.20	5 (1%) 77 75	62, 87, 122, 143	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	75	ILE	3.3
1	B	86	THR	3.0
1	A	47	ILE	2.4
1	B	47	ILE	2.2
1	A	73	LEU	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 monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.