



wwPDB X-ray Structure Validation Summary Report ⓘ

May 10, 2021 – 10:10 AM JST

PDB ID : 7DFT
Title : Crystal structure of Xanthomonas oryzae ClpP
Authors : Yang, C.-G.; Yang, T.
Deposited on : 2020-11-09
Resolution : 1.80 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 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	:	2.18
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)	:	2.18

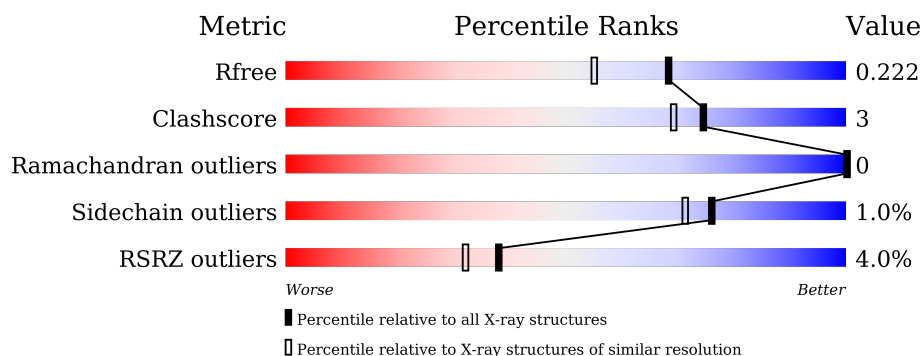
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	208	<div> <div>2%</div> <div> <div></div> <div>79%</div> <div>7%</div> <div>13%</div> </div> </div>
1	B	208	<div> <div>5%</div> <div> <div></div> <div>84%</div> <div>6%</div> <div>10%</div> </div> </div>
1	C	208	<div> <div>4%</div> <div> <div></div> <div>80%</div> <div>7%</div> <div>12%</div> </div> </div>
1	D	208	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>6%</div> <div>10%</div> </div> </div>
1	E	208	<div> <div>4%</div> <div> <div></div> <div>85%</div> <div>5%</div> <div>10%</div> </div> </div>
1	F	208	<div> <div>3%</div> <div> <div></div> <div>79%</div> <div>7%</div> <div>13%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	208	

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
2	CL	D	301	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10767 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 ATP-dependent Clp protease proteolytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	180	Total	C	N	O	S	0	1	0
			1376	870	236	263	7			
1	B	187	Total	C	N	O	S	0	0	0
			1424	896	249	272	7			
1	C	182	Total	C	N	O	S	0	0	0
			1382	872	239	264	7			
1	D	188	Total	C	N	O	S	0	0	0
			1429	900	251	271	7			
1	E	187	Total	C	N	O	S	0	0	0
			1432	901	250	274	7			
1	F	180	Total	C	N	O	S	0	0	0
			1369	864	237	262	6			
1	G	183	Total	C	N	O	S	0	0	0
			1392	880	242	263	7			

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	3	Total	Cl	0	0
			3	3		
2	B	1	Total	Cl	0	0
			1	1		
2	C	1	Total	Cl	0	0
			1	1		
2	D	1	Total	Cl	0	0
			1	1		
2	E	1	Total	Cl	0	0
			1	1		
2	G	1	Total	Cl	0	0
			1	1		

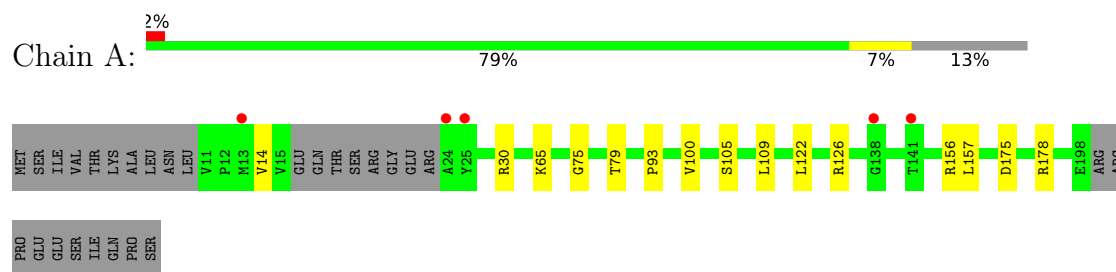
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	125	Total 125	O 125	0	0
3	B	141	Total 141	O 141	0	0
3	C	140	Total 140	O 140	0	0
3	D	139	Total 139	O 139	0	0
3	E	126	Total 126	O 126	0	0
3	F	133	Total 133	O 133	0	0
3	G	151	Total 151	O 151	0	0

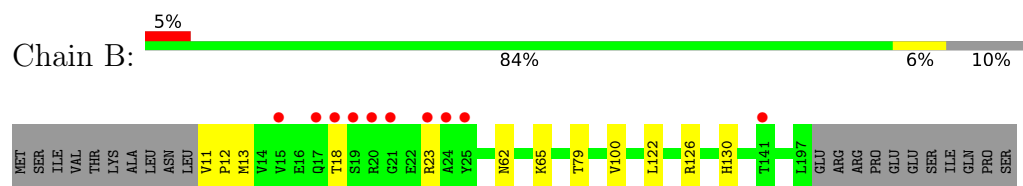
3 Residue-property plots [i](#)

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.

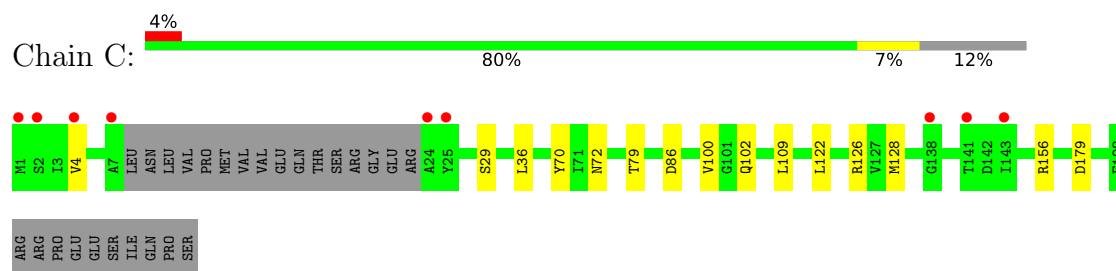
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



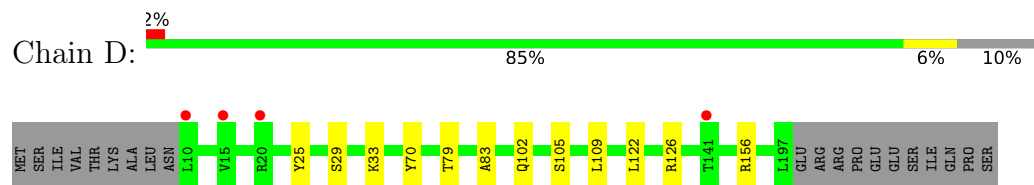
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



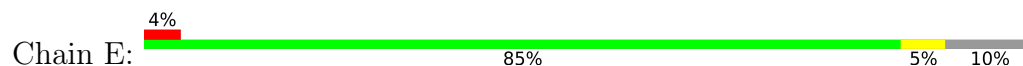
- Molecule 1: ATP-dependent Clp protease proteolytic subunit



- Molecule 1: ATP-dependent Clp protease proteolytic subunit

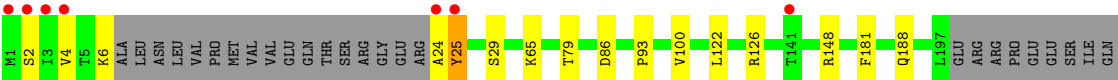
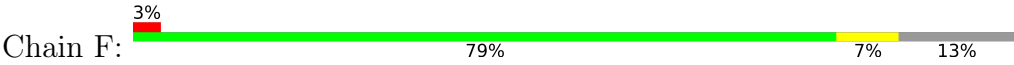


- Molecule 1: ATP-dependent Clp protease proteolytic subunit

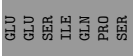
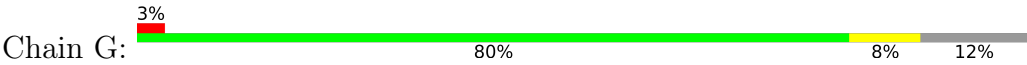




● Molecule 1: ATP-dependent Clp protease proteolytic subunit



● Molecule 1: ATP-dependent Clp protease proteolytic subunit



4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α , β , γ	132.55Å 132.55Å 187.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.64 – 1.80 29.64 – 1.79	Depositor EDS
% Data completeness (in resolution range)	92.6 (29.64-1.80) 92.6 (29.64-1.79)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.88 (at 1.79Å)	Xtriage
Refinement program	PHENIX 1.12_2829	Depositor
R, R_{free}	0.186 , 0.222 0.187 , 0.222	Depositor DCC
R_{free} test set	7270 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	15.0	Xtriage
Anisotropy	0.137	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 54.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10767	wwPDB-VP
Average B, all atoms (Å ²)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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](#)

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

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	0.41	0/1399	0.57	0/1896
1	B	0.46	0/1445	0.56	0/1960
1	C	0.47	0/1401	0.60	0/1897
1	D	0.49	0/1450	0.60	0/1966
1	E	0.44	0/1453	0.57	0/1969
1	F	0.48	0/1388	0.57	0/1880
1	G	0.50	0/1412	0.58	0/1914
All	All	0.46	0/9948	0.58	0/13482

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	1376	0	1390	10	0
1	B	1424	0	1433	8	0
1	C	1382	0	1398	17	0
1	D	1429	0	1442	10	0
1	E	1432	0	1448	11	0
1	F	1369	0	1384	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	1392	0	1405	12	0
2	A	3	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	2	0
2	E	1	0	0	1	0
2	G	1	0	0	0	0
3	A	125	0	0	0	0
3	B	141	0	0	1	0
3	C	140	0	0	1	0
3	D	139	0	0	0	0
3	E	126	0	0	2	0
3	F	133	0	0	1	0
3	G	151	0	0	1	0
All	All	10767	0	9900	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 3.

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:36:LEU:HD11	1:C:70:TYR:CD1	2.15	0.82
2:E:301:CL:CL	3:E:516:HOH:O	2.36	0.79
1:C:36:LEU:HD11	1:C:70:TYR:HD1	1.50	0.74
1:A:156:ARG:HH21	1:G:124:ASN:HD21	1.38	0.72
1:D:102:GLN:HE22	1:D:126:ARG:HH21	1.41	0.67

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	177/208 (85%)	175 (99%)	2 (1%)	0	100	100
1	B	185/208 (89%)	182 (98%)	3 (2%)	0	100	100
1	C	178/208 (86%)	173 (97%)	5 (3%)	0	100	100
1	D	186/208 (89%)	182 (98%)	4 (2%)	0	100	100
1	E	185/208 (89%)	181 (98%)	4 (2%)	0	100	100
1	F	176/208 (85%)	172 (98%)	4 (2%)	0	100	100
1	G	179/208 (86%)	176 (98%)	3 (2%)	0	100	100
All	All	1266/1456 (87%)	1241 (98%)	25 (2%)	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	147/173 (85%)	147 (100%)	0	100	100
1	B	151/173 (87%)	149 (99%)	2 (1%)	69	62
1	C	146/173 (84%)	144 (99%)	2 (1%)	67	59
1	D	151/173 (87%)	148 (98%)	3 (2%)	55	44
1	E	153/173 (88%)	153 (100%)	0	100	100
1	F	145/173 (84%)	143 (99%)	2 (1%)	67	59
1	G	147/173 (85%)	146 (99%)	1 (1%)	84	81
All	All	1040/1211 (86%)	1030 (99%)	10 (1%)	76	71

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	25	TYR
1	F	188	GLN
1	G	105	SER
1	C	109	LEU
1	D	105	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	130	HIS
1	G	124	ASN
1	D	102	GLN
1	C	158	ASN
1	F	188	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 8 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 ⓘ

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, 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	180/208 (86%)	-0.29	5 (2%) 53 47	9, 16, 39, 67	0
1	B	187/208 (89%)	-0.18	10 (5%) 26 21	9, 18, 48, 90	0
1	C	182/208 (87%)	-0.21	9 (4%) 29 24	8, 15, 41, 69	0
1	D	188/208 (90%)	-0.34	4 (2%) 63 59	8, 17, 39, 70	0
1	E	187/208 (89%)	-0.15	9 (4%) 30 25	10, 18, 44, 59	0
1	F	180/208 (86%)	-0.29	7 (3%) 39 33	10, 17, 37, 55	0
1	G	183/208 (87%)	-0.38	7 (3%) 40 35	8, 14, 39, 65	0
All	All	1287/1456 (88%)	-0.26	51 (3%) 38 32	8, 17, 41, 90	0

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	21	GLY	8.6
1	F	24	ALA	7.3
1	C	141	THR	6.1
1	B	18	THR	6.0
1	B	23	ARG	5.0

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 monosaccharides 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, 95th 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(Å ²)	Q<0.9
2	CL	D	301	1/1	0.92	0.06	53,53,53,53	0
2	CL	A	302	1/1	0.99	0.06	18,18,18,18	0
2	CL	A	303	1/1	0.99	0.04	25,25,25,25	0
2	CL	B	301	1/1	0.99	0.04	23,23,23,23	0
2	CL	C	301	1/1	0.99	0.05	18,18,18,18	0
2	CL	A	301	1/1	0.99	0.04	16,16,16,16	0
2	CL	E	301	1/1	0.99	0.04	24,24,24,24	0
2	CL	G	301	1/1	0.99	0.04	18,18,18,18	0

6.5 Other polymers [i](#)

There are no such residues in this entry.