



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:32 AM GMT

PDB ID : 2DF5
Title : Crystal Structure of Pf-PCP(1-204)-C
Authors : Katagiri, Y.; Takano, K.; Chon, H.; Matsumura, H.; Koga, Y.; Kanaya, S.
Deposited on : 2006-02-24
Resolution : 2.30 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
<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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

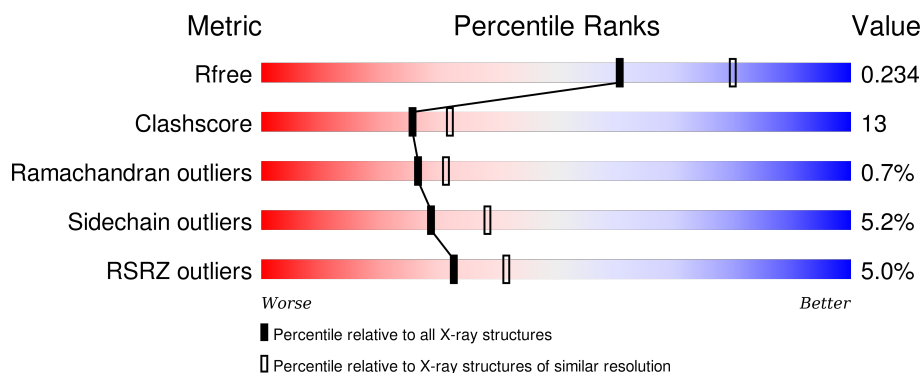
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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 ≥ 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	213	<div> <div>2%</div> <div>73%</div> <div>24%</div> <div>.</div> </div>
1	B	213	<div> <div>5%</div> <div>79%</div> <div>19%</div> <div>.</div> </div>
1	C	213	<div> <div>7%</div> <div>71%</div> <div>26%</div> <div>.</div> </div>
1	D	213	<div> <div>7%</div> <div>69%</div> <div>26%</div> <div>5%</div> </div>

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 7057 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 Pyrrolidone-carboxylate peptidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	213	Total	C	N	O	S	0	0	0
			1641	1059	271	302	9			
1	B	213	Total	C	N	O	S	0	0	0
			1641	1059	271	302	9			
1	C	213	Total	C	N	O	S	0	0	0
			1641	1059	271	302	9			
1	D	213	Total	C	N	O	S	0	0	0
			1641	1059	271	302	9			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	205	THR	-	SEE REMARK 999	UNP O73944
A	206	GLN	-	SEE REMARK 999	UNP O73944
A	207	ASP	-	SEE REMARK 999	UNP O73944
A	208	MET	-	SEE REMARK 999	UNP O73944
A	209	ILE	-	SEE REMARK 999	UNP O73944
A	210	ASN	-	SEE REMARK 999	UNP O73944
A	211	LYS	-	SEE REMARK 999	UNP O73944
A	212	SER	-	SEE REMARK 999	UNP O73944
A	213	THR	-	SEE REMARK 999	UNP O73944
B	205	THR	-	SEE REMARK 999	UNP O73944
B	206	GLN	-	SEE REMARK 999	UNP O73944
B	207	ASP	-	SEE REMARK 999	UNP O73944
B	208	MET	-	SEE REMARK 999	UNP O73944
B	209	ILE	-	SEE REMARK 999	UNP O73944
B	210	ASN	-	SEE REMARK 999	UNP O73944
B	211	LYS	-	SEE REMARK 999	UNP O73944
B	212	SER	-	SEE REMARK 999	UNP O73944
B	213	THR	-	SEE REMARK 999	UNP O73944
C	205	THR	-	SEE REMARK 999	UNP O73944
C	206	GLN	-	SEE REMARK 999	UNP O73944
C	207	ASP	-	SEE REMARK 999	UNP O73944

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	208	MET	-	SEE REMARK 999	UNP O73944
C	209	ILE	-	SEE REMARK 999	UNP O73944
C	210	ASN	-	SEE REMARK 999	UNP O73944
C	211	LYS	-	SEE REMARK 999	UNP O73944
C	212	SER	-	SEE REMARK 999	UNP O73944
C	213	THR	-	SEE REMARK 999	UNP O73944
D	205	THR	-	SEE REMARK 999	UNP O73944
D	206	GLN	-	SEE REMARK 999	UNP O73944
D	207	ASP	-	SEE REMARK 999	UNP O73944
D	208	MET	-	SEE REMARK 999	UNP O73944
D	209	ILE	-	SEE REMARK 999	UNP O73944
D	210	ASN	-	SEE REMARK 999	UNP O73944
D	211	LYS	-	SEE REMARK 999	UNP O73944
D	212	SER	-	SEE REMARK 999	UNP O73944
D	213	THR	-	SEE REMARK 999	UNP O73944

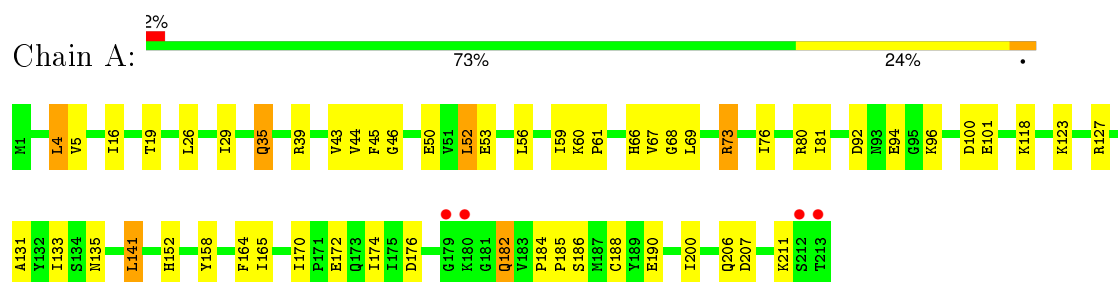
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	146	Total O 146 146	0	0
2	B	128	Total O 128 128	0	0
2	C	114	Total O 114 114	0	0
2	D	105	Total O 105 105	0	0

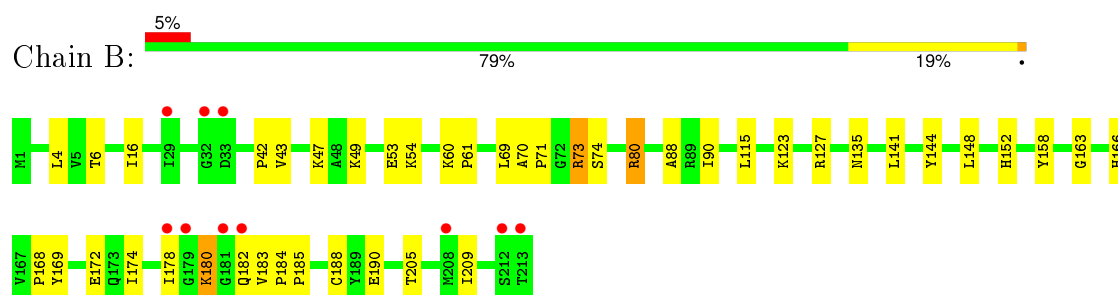
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

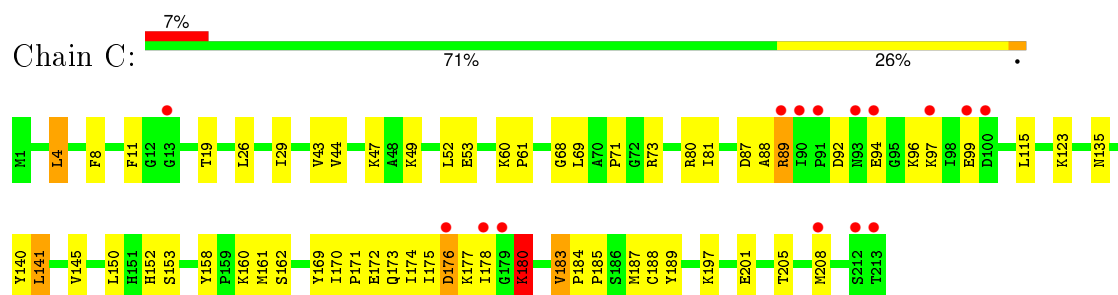
• Molecule 1: Pyrrolidone-carboxylate peptidase



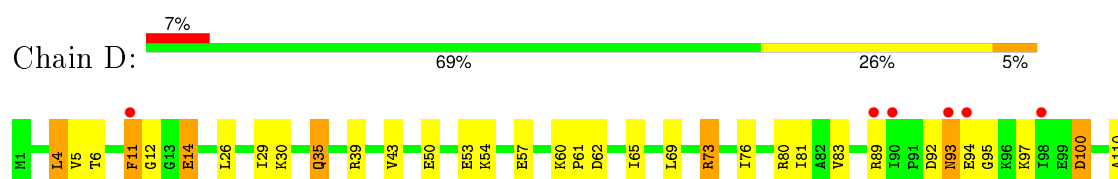
• Molecule 1: Pyrrolidone-carboxylate peptidase



• Molecule 1: Pyrrolidone-carboxylate peptidase



• Molecule 1: Pyrrolidone-carboxylate peptidase





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	104.24Å 104.45Å 48.25Å 90.00° 100.39° 90.00°	Depositor
Resolution (Å)	29.24 – 2.30 47.46 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.9 (29.24-2.30) 98.0 (47.46-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.89 (at 2.32Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.197 , 0.234 0.197 , 0.234	Depositor DCC
R_{free} test set	2200 reflections (5.00%)	DCC
Wilson B-factor (Å ²)	23.1	Xtriage
Anisotropy	0.676	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 54.6	EDS
Estimated twinning fraction	0.027 for -h-l,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 44026 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7057	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.08% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.34	0/1675	0.60	0/2264
1	B	0.33	0/1675	0.57	0/2264
1	C	0.32	0/1675	0.56	0/2264
1	D	0.31	0/1675	0.57	0/2264
All	All	0.33	0/6700	0.58	0/9056

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	1641	0	1707	41	0
1	B	1641	0	1707	40	0
1	C	1641	0	1707	51	0
1	D	1641	0	1707	43	0
2	A	146	0	0	6	0
2	B	128	0	0	7	0
2	C	114	0	0	6	0
2	D	105	0	0	4	0
All	All	7057	0	6828	168	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 13.

The worst 5 of 168 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:174:ILE:HD11	1:A:186:SER:HB3	1.58	0.85
1:C:80:ARG:HH11	1:C:135:ASN:HD21	1.28	0.80
1:A:188:CYS:HB3	2:C:311:HOH:O	1.82	0.79
1:C:170:ILE:HG22	1:C:173:GLN:HG3	1.67	0.77
1:B:180:LYS:H	1:B:180:LYS:HD2	1.50	0.76

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	211/213 (99%)	202 (96%)	9 (4%)	0	100	100
1	B	211/213 (99%)	202 (96%)	9 (4%)	0	100	100
1	C	211/213 (99%)	199 (94%)	10 (5%)	2 (1%)	21	24
1	D	211/213 (99%)	195 (92%)	12 (6%)	4 (2%)	10	8
All	All	844/852 (99%)	798 (94%)	40 (5%)	6 (1%)	26	31

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	180	LYS
1	D	11	PHE
1	D	209	ILE
1	D	100	ASP
1	D	95	GLY

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	179/179 (100%)	169 (94%)	10 (6%)	26	35
1	B	179/179 (100%)	175 (98%)	4 (2%)	60	77
1	C	179/179 (100%)	170 (95%)	9 (5%)	30	41
1	D	179/179 (100%)	165 (92%)	14 (8%)	16	19
All	All	716/716 (100%)	679 (95%)	37 (5%)	29	38

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	69	LEU
1	C	176	ASP
1	D	208	MET
1	C	73	ARG
1	C	89	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	152	HIS
1	C	93	ASN
1	D	35	GLN
1	B	135	ASN
1	C	152	HIS

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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	213/213 (100%)	0.06	4 (1%) 70 76	14, 21, 47, 83	0
1	B	213/213 (100%)	0.10	10 (4%) 35 44	13, 23, 59, 83	0
1	C	213/213 (100%)	0.30	15 (7%) 19 27	15, 25, 72, 88	0
1	D	213/213 (100%)	0.40	14 (6%) 22 29	17, 27, 73, 90	0
All	All	852/852 (100%)	0.22	43 (5%) 32 41	13, 25, 69, 90	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	213	THR	10.3
1	D	213	THR	7.4
1	D	90	ILE	6.7
1	C	178	ILE	6.6
1	A	213	THR	6.2

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.