



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 14, 2023 – 11:43 PM EDT

PDB ID : 7U1J
Title : Crystal structure of Pisum sativum convicilin
Authors : Bakestani, I.D.; Robinson, K.A.; Loewen, M.C.
Deposited on : 2022-02-21
Resolution : 2.70 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
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

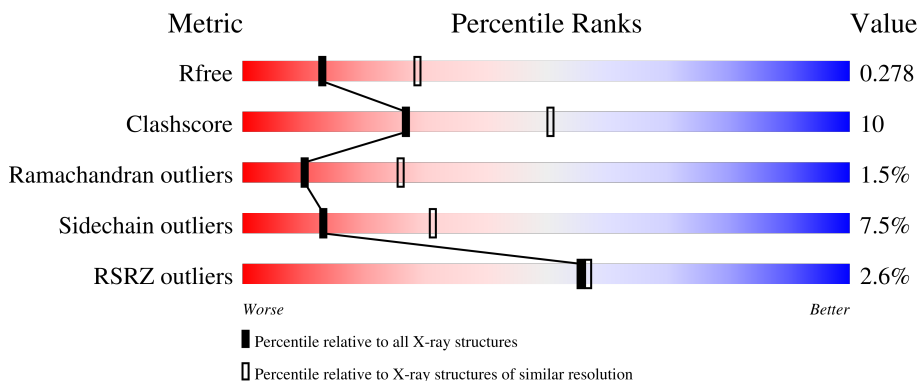
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	413	
1	B	413	
1	C	413	

2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 16811 atoms, of which 8362 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 Convicilin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	349	5622	1782	2820	484	534	2	59	0	0
1	B	344	5521	1754	2763	476	526	2	59	0	0
1	C	346	5548	1762	2779	478	527	2	59	0	0

There are 57 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	177	MET	-	expression tag	UNP Q9M3X6
A	178	GLY	-	expression tag	UNP Q9M3X6
A	179	SER	-	expression tag	UNP Q9M3X6
A	180	SER	-	expression tag	UNP Q9M3X6
A	181	HIS	-	expression tag	UNP Q9M3X6
A	182	HIS	-	expression tag	UNP Q9M3X6
A	183	HIS	-	expression tag	UNP Q9M3X6
A	184	HIS	-	expression tag	UNP Q9M3X6
A	185	HIS	-	expression tag	UNP Q9M3X6
A	186	HIS	-	expression tag	UNP Q9M3X6
A	187	LEU	-	expression tag	UNP Q9M3X6
A	188	VAL	-	expression tag	UNP Q9M3X6
A	189	PRO	-	expression tag	UNP Q9M3X6
A	190	ARG	-	expression tag	UNP Q9M3X6
A	191	GLY	-	expression tag	UNP Q9M3X6
A	192	SER	-	expression tag	UNP Q9M3X6
A	193	HIS	-	expression tag	UNP Q9M3X6
A	194	MET	-	expression tag	UNP Q9M3X6
A	195	MET	-	expression tag	UNP Q9M3X6
B	177	MET	-	expression tag	UNP Q9M3X6
B	178	GLY	-	expression tag	UNP Q9M3X6
B	179	SER	-	expression tag	UNP Q9M3X6
B	180	SER	-	expression tag	UNP Q9M3X6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	181	HIS	-	expression tag	UNP Q9M3X6
B	182	HIS	-	expression tag	UNP Q9M3X6
B	183	HIS	-	expression tag	UNP Q9M3X6
B	184	HIS	-	expression tag	UNP Q9M3X6
B	185	HIS	-	expression tag	UNP Q9M3X6
B	186	HIS	-	expression tag	UNP Q9M3X6
B	187	LEU	-	expression tag	UNP Q9M3X6
B	188	VAL	-	expression tag	UNP Q9M3X6
B	189	PRO	-	expression tag	UNP Q9M3X6
B	190	ARG	-	expression tag	UNP Q9M3X6
B	191	GLY	-	expression tag	UNP Q9M3X6
B	192	SER	-	expression tag	UNP Q9M3X6
B	193	HIS	-	expression tag	UNP Q9M3X6
B	194	MET	-	expression tag	UNP Q9M3X6
B	195	MET	-	expression tag	UNP Q9M3X6
C	177	MET	-	expression tag	UNP Q9M3X6
C	178	GLY	-	expression tag	UNP Q9M3X6
C	179	SER	-	expression tag	UNP Q9M3X6
C	180	SER	-	expression tag	UNP Q9M3X6
C	181	HIS	-	expression tag	UNP Q9M3X6
C	182	HIS	-	expression tag	UNP Q9M3X6
C	183	HIS	-	expression tag	UNP Q9M3X6
C	184	HIS	-	expression tag	UNP Q9M3X6
C	185	HIS	-	expression tag	UNP Q9M3X6
C	186	HIS	-	expression tag	UNP Q9M3X6
C	187	LEU	-	expression tag	UNP Q9M3X6
C	188	VAL	-	expression tag	UNP Q9M3X6
C	189	PRO	-	expression tag	UNP Q9M3X6
C	190	ARG	-	expression tag	UNP Q9M3X6
C	191	GLY	-	expression tag	UNP Q9M3X6
C	192	SER	-	expression tag	UNP Q9M3X6
C	193	HIS	-	expression tag	UNP Q9M3X6
C	194	MET	-	expression tag	UNP Q9M3X6
C	195	MET	-	expression tag	UNP Q9M3X6

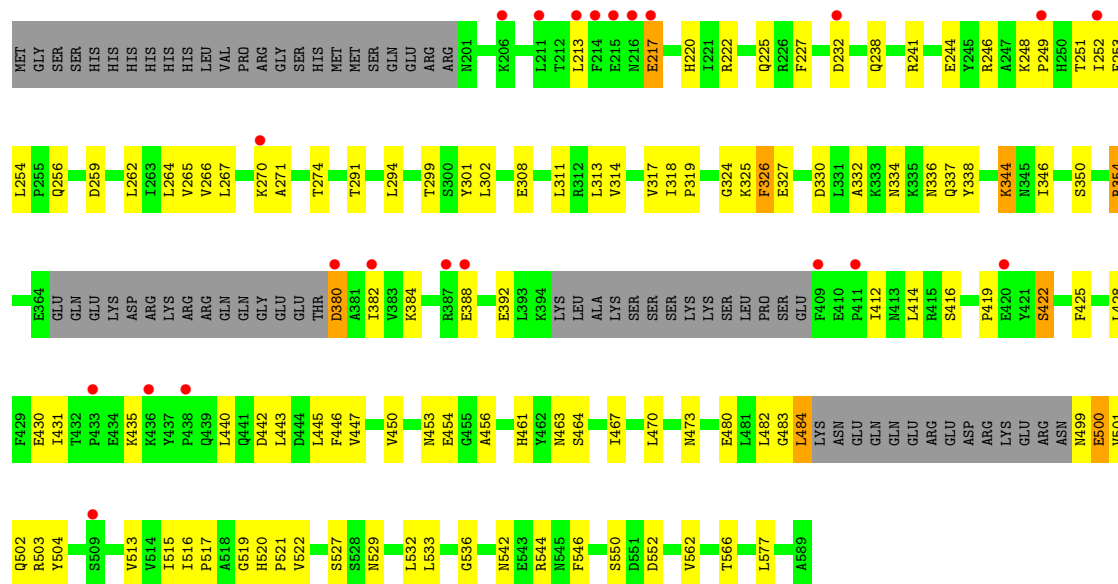
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	72	Total O 72 72	0	0
2	B	34	Total O 34 34	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	14	Total	O	0	0
			14	14		



4 Data and refinement statistics i

Property	Value	Source
Space group	F 2 3	Depositor
Cell constants a, b, c, α , β , γ	269.15Å 269.15Å 269.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	134.94 – 2.70 134.58 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.8 (134.94-2.70) 91.5 (134.58-2.70)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.89 (at 2.69Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.215 , 0.281 0.215 , 0.278	Depositor DCC
R_{free} test set	2069 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å ²)	40.3	Xtrriage
Anisotropy	0.000	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 39.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.027 for k,h,-l	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	16811	wwPDB-VP
Average B, all atoms (Å ²)	46.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.26% 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	0.74	0/2851	0.99	2/3855 (0.1%)
1	B	0.77	0/2807	0.99	2/3797 (0.1%)
1	C	0.79	0/2817	0.96	0/3811
All	All	0.77	0/8475	0.98	4/11463 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	C	0	1
All	All	0	2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	415	ARG	NE-CZ-NH1	5.66	123.13	120.30
1	A	226	ARG	CG-CD-NE	5.09	122.49	111.80
1	A	226	ARG	NE-CZ-NH1	5.09	122.84	120.30
1	B	551	ASP	CB-CA-C	5.03	120.45	110.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	232	ASP	Peptide
1	C	464	SER	Peptide

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	2802	2820	2809	41	1
1	B	2758	2763	2747	59	1
1	C	2769	2779	2760	69	0
2	A	72	0	0	1	0
2	B	34	0	0	1	0
2	C	14	0	0	1	0
All	All	8449	8362	8316	169	1

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 10.

The worst 5 of 169 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:562:VAL:O	1:C:566:THR:HG22	1.69	0.92
1:B:271:ALA:HB2	1:B:311:LEU:HD22	1.61	0.82
1:B:250:HIS:ND1	1:B:390:ILE:HD13	1.96	0.80
1:A:562:VAL:O	1:A:566:THR:HG22	1.82	0.78
1:B:423:ASN:HB3	1:B:425:PHE:H	1.55	0.69

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:387:ARG:HH21	1:B:573:GLU:OE1[48_445]	1.58	0.02

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	341/413 (83%)	323 (95%)	15 (4%)	3 (1%)	17	40
1	B	336/413 (81%)	307 (91%)	22 (6%)	7 (2%)	7	18
1	C	338/413 (82%)	304 (90%)	29 (9%)	5 (2%)	10	26
All	All	1015/1239 (82%)	934 (92%)	66 (6%)	15 (2%)	10	26

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	436	LYS
1	B	335	LYS
1	C	500	GLU
1	A	232	ASP
1	B	296	ALA

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	314/374 (84%)	293 (93%)	21 (7%)	16	37
1	B	307/374 (82%)	283 (92%)	24 (8%)	12	29
1	C	308/374 (82%)	283 (92%)	25 (8%)	11	27
All	All	929/1122 (83%)	859 (92%)	70 (8%)	13	31

5 of 70 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	380	ASP
1	C	388	GLU
1	C	473	ASN
1	B	283	SER
1	B	277	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	538	ASN
1	B	585	HIS
1	C	583	GLN
1	A	588	ASN
1	A	336	ASN

5.3.3 RNA [i](#)

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 monosaccharides 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	349/413 (84%)	0.06	0 100 100	16, 30, 56, 86	0
1	B	344/413 (83%)	0.06	5 (1%) 73 76	24, 41, 77, 111	0
1	C	346/413 (83%)	0.40	22 (6%) 19 18	31, 55, 96, 128	0
All	All	1039/1239 (83%)	0.18	27 (2%) 56 57	16, 42, 83, 128	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	214	PHE	4.1
1	C	388	GLU	4.1
1	C	409	PHE	3.4
1	C	216	ASN	3.4
1	B	335	LYS	3.2

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 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.