



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 06:03 PM UTC

PDB ID : 7U1I / pdb_00007u1i
Title : Crystal structure of Pisum sativum vicilin
Authors : Beavington, B.A.G.; Bakestani, I.D.; Robinson, K.A.; Loewen, M.C.
Deposited on : 2022-02-21
Resolution : 3.10 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

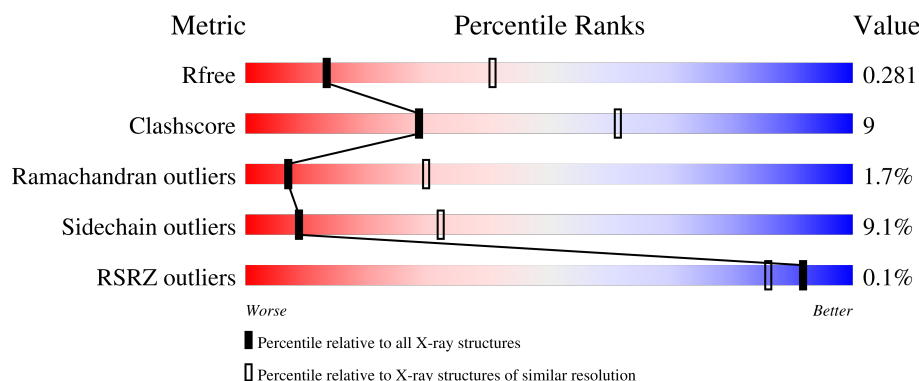
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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}	180053	1456 (3.10-3.10)
Clashscore	190562	1539 (3.10-3.10)
Ramachandran outliers	187476	1467 (3.10-3.10)
Sidechain outliers	187428	1467 (3.10-3.10)
RSRZ outliers	180081	1456 (3.10-3.10)

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	428	
1	B	428	
1	C	428	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 16934 atoms, of which 8426 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 Vicilin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	362	Total	C	H	N	O	61	0	0
			5708	1812	2850	490	556			
1	B	360	Total	C	H	N	O	60	0	0
			5666	1801	2824	488	553			
1	C	359	Total	C	H	N	O	59	0	0
			5540	1770	2752	475	543			

There are 57 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-17	MET	-	expression tag	UNP Q702P1
A	-16	GLY	-	expression tag	UNP Q702P1
A	-15	SER	-	expression tag	UNP Q702P1
A	-14	SER	-	expression tag	UNP Q702P1
A	-13	HIS	-	expression tag	UNP Q702P1
A	-12	HIS	-	expression tag	UNP Q702P1
A	-11	HIS	-	expression tag	UNP Q702P1
A	-10	HIS	-	expression tag	UNP Q702P1
A	-9	HIS	-	expression tag	UNP Q702P1
A	-8	HIS	-	expression tag	UNP Q702P1
A	-7	LEU	-	expression tag	UNP Q702P1
A	-6	VAL	-	expression tag	UNP Q702P1
A	-5	PRO	-	expression tag	UNP Q702P1
A	-4	ARG	-	expression tag	UNP Q702P1
A	-3	GLY	-	expression tag	UNP Q702P1
A	-2	SER	-	expression tag	UNP Q702P1
A	-1	HIS	-	expression tag	UNP Q702P1
A	0	MET	-	expression tag	UNP Q702P1
A	1	MET	-	expression tag	UNP Q702P1
B	-17	MET	-	expression tag	UNP Q702P1
B	-16	GLY	-	expression tag	UNP Q702P1
B	-15	SER	-	expression tag	UNP Q702P1
B	-14	SER	-	expression tag	UNP Q702P1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-13	HIS	-	expression tag	UNP Q702P1
B	-12	HIS	-	expression tag	UNP Q702P1
B	-11	HIS	-	expression tag	UNP Q702P1
B	-10	HIS	-	expression tag	UNP Q702P1
B	-9	HIS	-	expression tag	UNP Q702P1
B	-8	HIS	-	expression tag	UNP Q702P1
B	-7	LEU	-	expression tag	UNP Q702P1
B	-6	VAL	-	expression tag	UNP Q702P1
B	-5	PRO	-	expression tag	UNP Q702P1
B	-4	ARG	-	expression tag	UNP Q702P1
B	-3	GLY	-	expression tag	UNP Q702P1
B	-2	SER	-	expression tag	UNP Q702P1
B	-1	HIS	-	expression tag	UNP Q702P1
B	0	MET	-	expression tag	UNP Q702P1
B	1	MET	-	expression tag	UNP Q702P1
C	-17	MET	-	expression tag	UNP Q702P1
C	-16	GLY	-	expression tag	UNP Q702P1
C	-15	SER	-	expression tag	UNP Q702P1
C	-14	SER	-	expression tag	UNP Q702P1
C	-13	HIS	-	expression tag	UNP Q702P1
C	-12	HIS	-	expression tag	UNP Q702P1
C	-11	HIS	-	expression tag	UNP Q702P1
C	-10	HIS	-	expression tag	UNP Q702P1
C	-9	HIS	-	expression tag	UNP Q702P1
C	-8	HIS	-	expression tag	UNP Q702P1
C	-7	LEU	-	expression tag	UNP Q702P1
C	-6	VAL	-	expression tag	UNP Q702P1
C	-5	PRO	-	expression tag	UNP Q702P1
C	-4	ARG	-	expression tag	UNP Q702P1
C	-3	GLY	-	expression tag	UNP Q702P1
C	-2	SER	-	expression tag	UNP Q702P1
C	-1	HIS	-	expression tag	UNP Q702P1
C	0	MET	-	expression tag	UNP Q702P1
C	1	MET	-	expression tag	UNP Q702P1

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	9	Total O 9 9	0	0
2	B	9	Total O 9 9	0	0

Continued on next page...

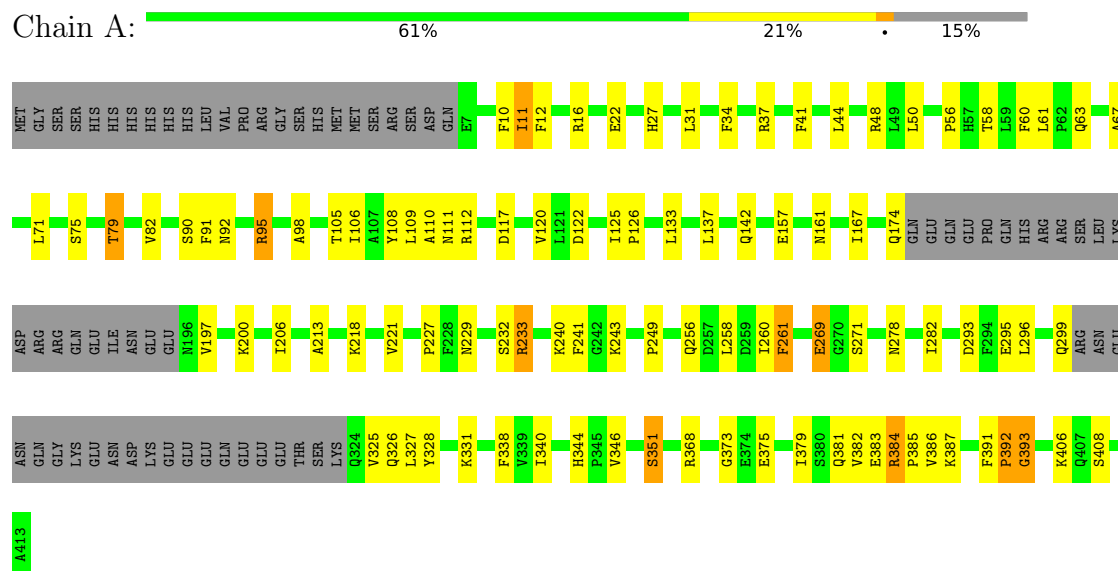
Continued from previous page...

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

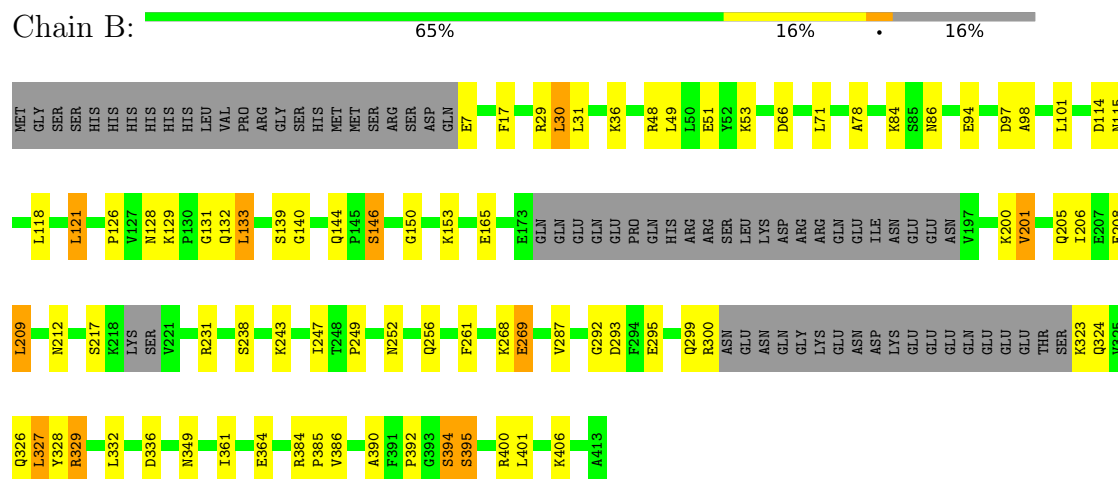
3 Residue-property plots

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: Vicilin



• Molecule 1: Vicilin



• Molecule 1: Vicilin



GLU	GLU	GLU	THR	S322	K323	Q324	V325	Q326	L327	Y328	S333	P334	V339	I340	P341	P345	V346	A347	I348	N349	F359	R368	N369	F370	V378	Q381	V382	E383	R384	P385	E388	K403	N404	Q405	K406	Q407	A411	N412	A413																
VAL	S222	S223	E224	P227	R231	K243	E246	P249	E250	K251	Q253	Q254	L255	Q256	D259	I260	F261	V262	N263	D266	E269	G270	S271	L272	L273	N276	Q290	E295	L296	V297	G298	Q299	ARG	ASN	GLU	ASN	GLN	GLY	LYS	GLU	GLU	GLN													
L147	L148	S149	G150	F151	S152	K153	N154	I155	L156	E157	N163	Y164	E165	E166	E173	GLN	GLU	GLN	GLN	PRO	GLN	HIS	ARG	ARG	SER	LEU	LYS	ASP	ARG	ARG	ARG	GLN	GLU	ILE	ASN	GLU	GLU	ASN	V197	I198	V199	K200	V201	I206	E207	E208	L209	S210	K211	N212	S216	S217	LYS	LYS	SER
L61	P62	A67	D68	F69	I70	L71	V72	V73	T79	L80	T81	V82	L83	K84	R88	F91	N92	L93	D97	A98	I99	K100	L101	Y108	L109	A110	N111	D114	N115	L118	L121	D122	L123	A124	I125	P126	V127	N128	K129	P130	G131	Q132	L133	G140	T141	Q144	P145	L146							
MET	GLY	SER	HIS	HIS	HIS	HIS	HIS	HIS	LEU	VAL	PRO	ARG	GLY	SER	MET	MET	SER	ARG	SER	ASP	Q6	E7	N8	I11	Q18	N23	E24	N25	R29	L30	I31	Q32	K33	K36	E42	N43	Q45	N46	Y47	R48	L49	L50	E51	K55	P56	H57	T58	L59	F60						

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	98.32Å 52.69Å 127.84Å 90.00° 95.19° 90.00°	Depositor
Resolution (Å)	98.10 – 3.10 98.10 – 3.10	Depositor EDS
% Data completeness (in resolution range)	84.4 (98.10-3.10) 68.2 (98.10-3.10)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.81 (at 3.13Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.185 , 0.285 0.181 , 0.281	Depositor DCC
R_{free} test set	1022 reflections (2.23%)	wwPDB-VP
Wilson B-factor (Å ²)	50.6	Xtriage
Anisotropy	0.373	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 53.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	16934	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.74% 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	1.03	0/2907	1.40	4/3930 (0.1%)
1	B	1.04	0/2890	1.42	2/3908 (0.1%)
1	C	1.06	0/2836	1.39	0/3844
All	All	1.04	0/8633	1.40	6/11682 (0.1%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	261	PHE	CA-CB-CG	6.46	120.26	113.80
1	A	105	THR	CA-CB-OG1	-6.05	100.52	109.60
1	A	197	VAL	N-CA-C	-5.71	107.18	112.43
1	A	373	GLY	CA-C-O	-5.56	117.46	122.57
1	A	117	ASP	CA-CB-CG	5.27	117.87	112.60

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	2858	2850	2832	50	0
1	B	2842	2824	2804	41	0
1	C	2788	2752	2710	68	0
2	A	9	0	0	0	0
2	B	9	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	2	0	0	0	0
All	All	8508	8426	8346	148	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 9.

The worst 5 of 148 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:24:GLU:OE1	1:C:200:LYS:NZ	2.20	0.74
1:B:71:LEU:HD13	1:B:101:LEU:HD13	1.71	0.72
1:B:29:ARG:NH1	1:B:51:GLU:OE1	2.26	0.69
1:A:392:PRO:O	1:A:393:GLY:O	2.16	0.63
1:B:150:GLY:O	1:C:327:LEU:HB2	2.00	0.61

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	356/428 (83%)	310 (87%)	42 (12%)	4 (1%)	11	39
1	B	352/428 (82%)	303 (86%)	45 (13%)	4 (1%)	11	39
1	C	351/428 (82%)	301 (86%)	40 (11%)	10 (3%)	4	20
All	All	1059/1284 (82%)	914 (86%)	127 (12%)	18 (2%)	7	30

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	393	GLY
1	C	212	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	323	LYS
1	C	140	GLY
1	C	210	SER

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	318/386 (82%)	300 (94%)	18 (6%)	18	49
1	B	315/386 (82%)	285 (90%)	30 (10%)	8	30
1	C	303/386 (78%)	266 (88%)	37 (12%)	5	20
All	All	936/1158 (81%)	851 (91%)	85 (9%)	9	32

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	55	LYS
1	C	198	ILE
1	C	79	THR
1	C	115	ASN
1	C	212	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	366	ASN
1	B	377	ASN
1	C	404	ASN
1	C	92	ASN
1	B	32	GLN

5.3.3 RNA ⓘ

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 oligosaccharides 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	362/428 (84%)	-0.56	0	100	100	20, 40, 63, 92	0
1	B	360/428 (84%)	-0.49	0	100	100	20, 42, 84, 142	0
1	C	359/428 (83%)	-0.30	1 (0%)	90	80	35, 57, 80, 99	0
All	All	1081/1284 (84%)	-0.45	1 (0%)	92	86	20, 46, 78, 142	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	322	SER	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 oligosaccharides 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.