WORLDWIDE
wwPDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 28, 2025 – 09:22 PM EDT

PDB ID	:	7IH9 / pdb_00007ih9
Title	:	Endothiapepsin crystal D11a from ECBL-96 fragment library screening campaign used for PanDDA ground state calculation
Authors	:	Wollenhaupt, J.; Benz, L.S.; Jirgensons, A.; Miletic, T.; Mueller, U.; Weiss, M.S.
Deposited on	:	2025-06-05
Resolution	:	1.29 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

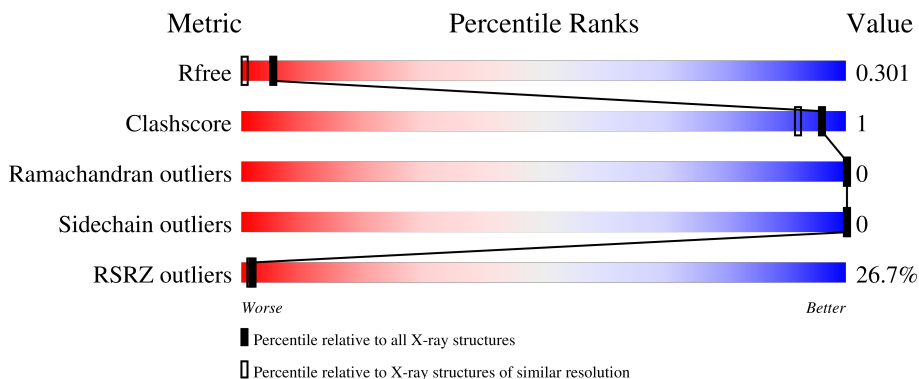|		
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.46

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 1.29 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	Percentile Ranks	Value
Rfree		0.301
Clashscore		1
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		26.7%

Worse   *Better*

▌Percentile relative to all X-ray structures
▯Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	164625	1387 (1.30-1.30)
Clashscore	180529	1497 (1.30-1.30)
Ramachandran outliers	177936	1455 (1.30-1.30)
Sidechain outliers	177891	1455 (1.30-1.30)
RSRZ outliers	164620	1384 (1.30-1.30)

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 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 <=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	330	27% 97% •

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 4927 atoms, of which 2341 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 Endothiapepsin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	330	Total 4803	C 1566	H 2341	N 367	O 527	S 2	0	25	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	124	Total 124	O 124	0	0

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: Endothiapepsin
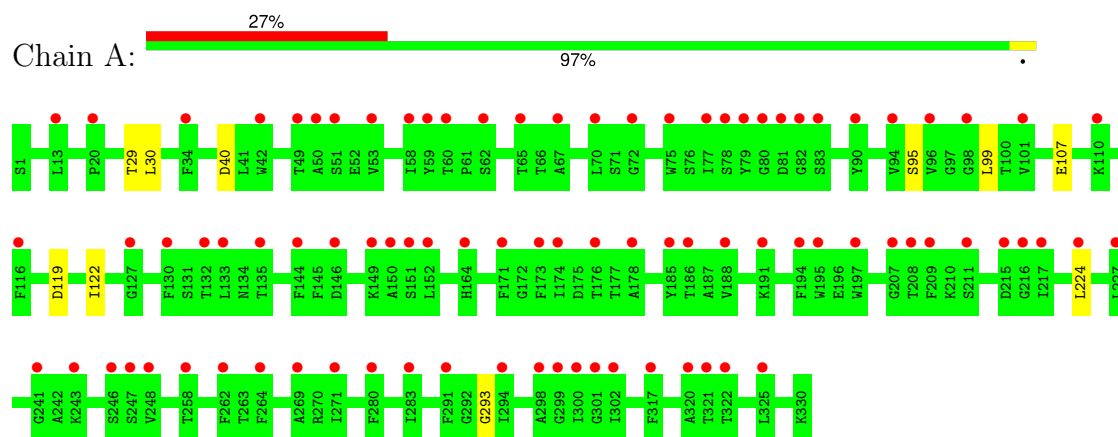
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	45.29Å 72.97Å 52.63Å 90.00° 109.39° 90.00°	Depositor
Resolution (Å)	49.64 – 1.29 49.64 – 1.29	Depositor EDS
% Data completeness (in resolution range)	92.1 (49.64-1.29) 92.3 (49.64-1.29)	Depositor EDS
R_{merge}	0.24	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.96 (at 1.29Å)	Xtriage
Refinement program	PHENIX (1.17_3644: ???)	Depositor
R, R_{free}	0.267 , 0.302 0.266 , 0.301	Depositor DCC
R_{free} test set	2055 reflections (2.74%)	wwPDB-VP
Wilson B-factor (Å²)	14.0	Xtriage
Anisotropy	0.135	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.44 , 31.8	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	4927	wwPDB-VP
Average B, all atoms (Å²)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.34% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.46	0/2594	0.63	0/3553

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	2462	2341	2300	5	0
2	A	124	0	0	0	0
All	All	2586	2341	2300	5	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 1.

All (5) 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:119:ASP:OD2	1:A:122:ILE:HD12	2.12	0.49
1:A:29:THR:C	1:A:30:LEU:HD12	2.38	0.49
1:A:224:LEU:HD22	1:A:293:GLY:HA2	1.97	0.46
1:A:95:SER:HA	1:A:99:LEU:O	2.18	0.44
1:A:40:ASP:OD2	1:A:107:GLU:OE2	2.39	0.41

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	353/330 (107%)	349 (99%)	4 (1%)	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	273/263 (104%)	273 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 ⓘ

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	330/330 (100%)	1.53	88 (26%)	2	1	8, 16, 26, 38	19 (5%)

All (88) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	300[A]	ILE	7.5
1	A	150[A]	ALA	5.8
1	A	321	THR	5.5
1	A	299[A]	GLY	5.1
1	A	246[A]	SER	4.7
1	A	248	VAL	4.7
1	A	174[A]	ILE	4.1
1	A	77	ILE	3.8
1	A	78	SER	3.6
1	A	302	ILE	3.6
1	A	50	ALA	3.5
1	A	258	THR	3.4
1	A	79	TYR	3.4
1	A	130	PHE	3.3
1	A	322	THR	3.3
1	A	271	ILE	3.3
1	A	188	VAL	3.3
1	A	194	PHE	3.3
1	A	320	ALA	3.3
1	A	298[A]	ALA	3.2
1	A	49	THR	3.2
1	A	208	THR	3.2
1	A	133	LEU	3.1
1	A	81	ASP	3.0
1	A	207	GLY	2.9
1	A	216	GLY	2.9
1	A	75	TRP	2.8

Continued on next page...

Full wwPDB X-ray Structure Validation Report 7IH9

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	301[A]	GLY	2.8
1	A	53	VAL	2.8
1	A	127	GLY	2.8
1	A	98	GLY	2.7
1	A	146	ASP	2.7
1	A	96	VAL	2.7
1	A	67	ALA	2.6
1	A	269	ALA	2.6
1	A	247	SER	2.6
1	A	171	PHE	2.6
1	A	173	PHE	2.6
1	A	51	SER	2.5
1	A	116	PHE	2.5
1	A	65	THR	2.5
1	A	90	TYR	2.5
1	A	291	PHE	2.4
1	A	283	ILE	2.4
1	A	83	SER	2.4
1	A	60	THR	2.4
1	A	132	THR	2.4
1	A	197	TRP	2.4
1	A	241	GLY	2.4
1	A	209	PHE	2.4
1	A	195	TRP	2.4
1	A	325[A]	LEU	2.3
1	A	149[A]	LYS	2.3
1	A	62	SER	2.3
1	A	217	ILE	2.3
1	A	59	TYR	2.3
1	A	151	SER	2.2
1	A	144	PHE	2.2
1	A	13	LEU	2.2
1	A	82	GLY	2.2
1	A	135	THR	2.2
1	A	243	LYS	2.2
1	A	42	TRP	2.2
1	A	80	GLY	2.2
1	A	110	LYS	2.2
1	A	215[A]	ASP	2.2
1	A	70	LEU	2.2
1	A	224	LEU	2.2
1	A	280	PHE	2.2

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	101	VAL	2.1
1	A	227	LEU	2.1
1	A	178	ALA	2.1
1	A	34	PHE	2.1
1	A	72	GLY	2.1
1	A	191	LYS	2.1
1	A	58	ILE	2.1
1	A	20	PRO	2.1
1	A	211	SER	2.1
1	A	185	TYR	2.1
1	A	294	ILE	2.1
1	A	164	HIS	2.1
1	A	152	LEU	2.0
1	A	262	PHE	2.0
1	A	264	PHE	2.0
1	A	317	PHE	2.0
1	A	94	VAL	2.0
1	A	176	THR	2.0
1	A	186	THR	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.