



# Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 02:17 am BST

PDB ID : 1FJU  
Title : THERMOLYSIN (80% ACETONITRILE SOAKED CRYSTALS)  
Authors : English, A.C.; Groom, C.R.; Hubbard, R.E.  
Deposited on : 2000-08-08  
Resolution : 2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

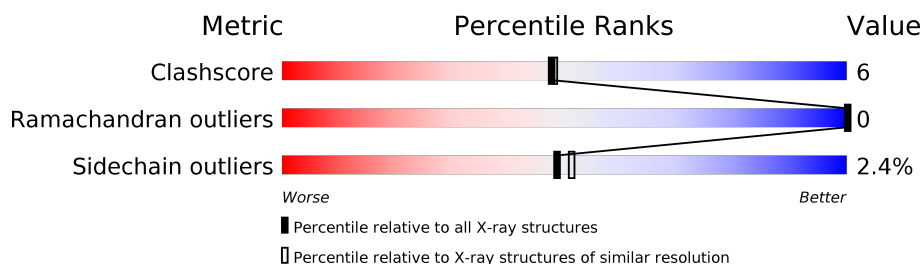
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	316	

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
5	CCN	A	507	-	-	X	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 2635 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 THERMOLYSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	316	Total	C	N	O	S	0	7	0
			2461	1544	415	499	3			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

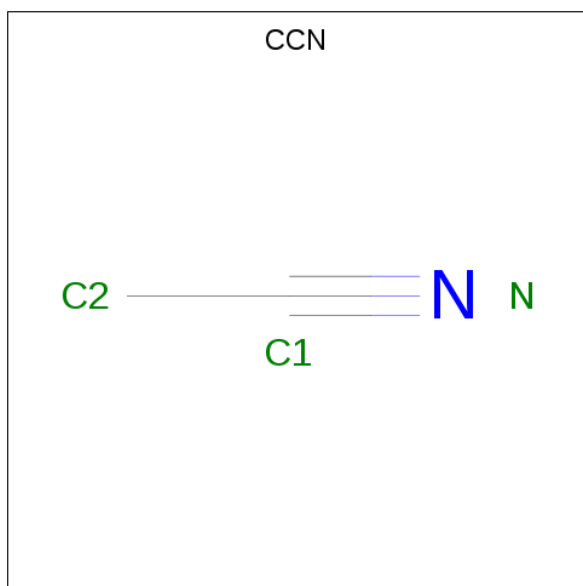
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total	Ca	0	0
			4	4		

- Molecule 4 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C<sub>2</sub>H<sub>6</sub>OS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	O	S	0	0
			4	2	1	1		

- Molecule 5 is ACETONITRILE (three-letter code: CCN) (formula:  $C_2H_3N$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	N	0	0
			3	2	1		

- Molecule 6 is water.

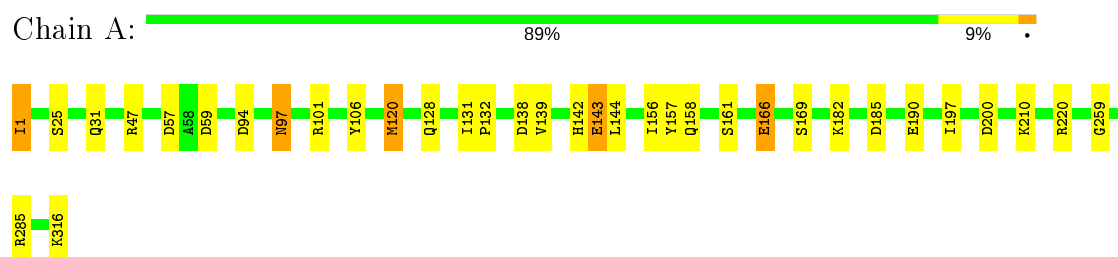
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	162	Total	O	0	0
			162	162		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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. 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. 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.

Note EDS was not executed.

#### • Molecule 1: THERMOLYSIN



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.86 Å 93.86 Å 131.04 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.00	Depositor
% Data completeness (in resolution range)	97.0 (20.00-2.00)	Depositor
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.154 , 0.205	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2635	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, CCN, DMS, CA

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	0.71	2/2557 (0.1%)	1.47	23/3478 (0.7%)

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	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	143[A]	GLU	C-N	-5.23	1.22	1.34
1	A	143[B]	GLU	C-N	-5.23	1.22	1.34

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	200	ASP	CB-CG-OD2	16.36	133.02	118.30
1	A	47	ARG	CD-NE-CZ	15.04	144.66	123.60
1	A	166	GLU	OE1-CD-OE2	-12.23	108.62	123.30
1	A	57	ASP	CB-CG-OD2	10.56	127.81	118.30
1	A	59	ASP	CB-CG-OD2	10.12	127.41	118.30
1	A	190	GLU	OE1-CD-OE2	-9.57	111.82	123.30
1	A	138	ASP	CB-CG-OD1	8.81	126.23	118.30
1	A	220	ARG	NE-CZ-NH2	-8.60	116.00	120.30
1	A	47	ARG	NE-CZ-NH2	-8.19	116.20	120.30
1	A	143[A]	GLU	OE1-CD-OE2	-8.00	113.70	123.30
1	A	143[B]	GLU	OE1-CD-OE2	-8.00	113.70	123.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	47	ARG	NE-CZ-NH1	7.98	124.29	120.30
1	A	185	ASP	CB-CG-OD1	7.54	125.08	118.30
1	A	200	ASP	CB-CG-OD1	-7.50	111.55	118.30
1	A	197	ILE	CA-C-N	7.04	132.69	117.20
1	A	285	ARG	NE-CZ-NH1	6.68	123.64	120.30
1	A	106	TYR	CB-CG-CD2	-6.11	117.33	121.00
1	A	120[A]	MET	CA-CB-CG	-5.79	103.46	113.30
1	A	120[B]	MET	CA-CB-CG	-5.79	103.46	113.30
1	A	94	ASP	CB-CG-OD2	-5.45	113.40	118.30
1	A	143[A]	GLU	CG-CD-OE2	5.29	128.89	118.30
1	A	143[B]	GLU	CG-CD-OE2	5.29	128.89	118.30
1	A	197	ILE	O-C-N	-5.09	114.55	122.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	25	SER	Mainchain
1	A	259	GLY	Mainchain

## 5.2 Too-close contacts ⓘ

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	2461	0	2296	28	0
2	A	1	0	0	0	0
3	A	4	0	0	0	0
4	A	4	0	6	0	0
5	A	3	0	3	2	0
6	A	162	0	0	2	0
All	All	2635	0	2305	29	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 6.

All (29) 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:120[B]:MET:CE	1:A:144:LEU:HG	1.40	1.48
1:A:120[B]:MET:HE3	1:A:144:LEU:CG	1.58	1.32
1:A:120[B]:MET:CE	1:A:144:LEU:CG	2.17	1.15
1:A:120[A]:MET:SD	1:A:143[A]:GLU:HB3	2.19	0.83
1:A:120[B]:MET:HE1	1:A:144:LEU:CG	2.13	0.79
1:A:120[B]:MET:CE	1:A:144:LEU:CD2	2.67	0.72
1:A:156:ILE:HD12	1:A:158:GLN:HG2	1.73	0.69
1:A:157[A]:TYR:OH	1:A:166:GLU:OE2	2.09	0.66
1:A:1:ILE:N	1:A:31:GLN:HE22	1.93	0.66
1:A:120[B]:MET:SD	1:A:144:LEU:CD2	2.85	0.64
1:A:120[B]:MET:CE	1:A:144:LEU:CD1	2.81	0.58
1:A:120[B]:MET:HE3	1:A:144:LEU:HG	0.64	0.56
1:A:139:VAL:O	1:A:143[B]:GLU:HG2	2.06	0.55
1:A:120[B]:MET:SD	1:A:144:LEU:HD23	2.48	0.54
1:A:101[B]:ARG:NH2	6:A:638:HOH:O	2.41	0.54
1:A:120[B]:MET:HE1	1:A:144:LEU:CD2	2.36	0.54
1:A:120[B]:MET:HE1	1:A:144:LEU:CD1	2.37	0.53
1:A:120[B]:MET:SD	1:A:144:LEU:HG	2.44	0.53
1:A:120[B]:MET:HE1	1:A:144:LEU:HD21	1.91	0.52
1:A:120[B]:MET:HE1	1:A:144:LEU:HD11	1.92	0.51
1:A:1:ILE:H2	1:A:31:GLN:HE22	1.57	0.50
5:A:507:CCN:H23	6:A:663:HOH:O	2.13	0.47
1:A:120[B]:MET:HE3	1:A:144:LEU:CD1	2.41	0.43
1:A:156:ILE:HG13	1:A:161[A]:SER:OG	2.18	0.43
1:A:131:ILE:HB	1:A:132:PRO:CD	2.49	0.42
1:A:142:HIS:CG	1:A:169:SER:HB3	2.54	0.41
1:A:143[B]:GLU:OE2	5:A:507:CCN:N	2.54	0.41
1:A:97:ASN:HA	1:A:97:ASN:HD22	1.69	0.41
1:A:1:ILE:H3	1:A:31:GLN:HE22	1.68	0.40

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	321/316 (102%)	311 (97%)	10 (3%)	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	259/252 (103%)	253 (98%)	6 (2%)	50	53

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	ILE
1	A	97	ASN
1	A	128	GLN
1	A	182	LYS
1	A	210	LYS
1	A	316	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	31	GLN
1	A	33	ASN
1	A	97	ASN
1	A	290	GLN
1	A	308	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 5 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	DMS	A	506	-	3,3,3	0.66	0	3,3,3	0.82	0
5	CCN	A	507	-	2,2,2	0.29	0	1,1,1	0.18	0

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.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	507	CCN	2	0

## 5.7 Other polymers [i](#)

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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.