



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 2, 2023 – 04:22 AM EDT

PDB ID : 3MH5
Title : HtrA proteases are activated by a conserved mechanism that can be triggered by distinct molecular cues
Authors : Krojer, T.; Sawa, J.; Huber, R.; Clausen, T.
Deposited on : 2010-04-07
Resolution : 3.00 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
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

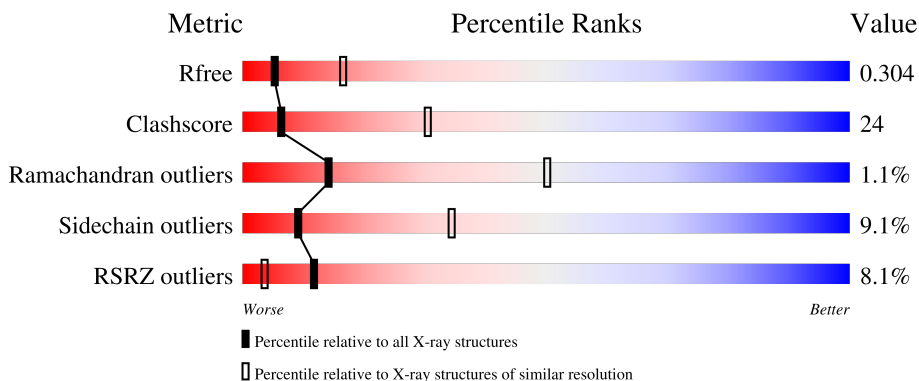
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	456	
1	B	456	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 2922 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.

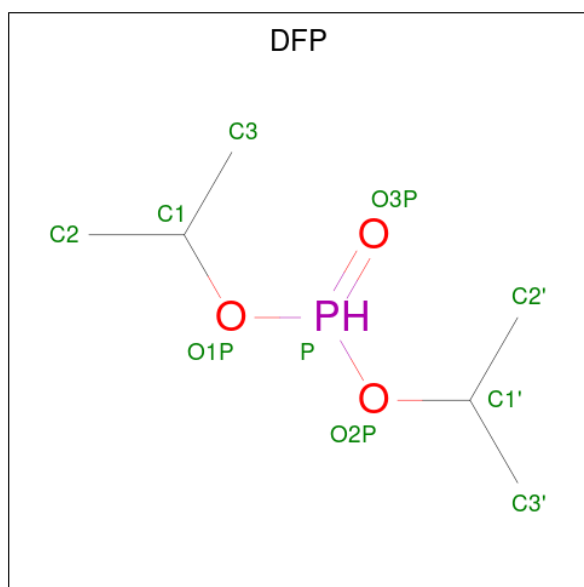
- Molecule 1 is a protein called Protease do.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	191	1409	883	241	277	8	0	0	0
1	B	202	1484	926	259	290	9	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	449	ARG	-	expression tag	UNP P0C0V0
A	450	SER	-	expression tag	UNP P0C0V0
A	451	HIS	-	expression tag	UNP P0C0V0
A	452	HIS	-	expression tag	UNP P0C0V0
A	453	HIS	-	expression tag	UNP P0C0V0
A	454	HIS	-	expression tag	UNP P0C0V0
A	455	HIS	-	expression tag	UNP P0C0V0
A	456	HIS	-	expression tag	UNP P0C0V0
B	449	ARG	-	expression tag	UNP P0C0V0
B	450	SER	-	expression tag	UNP P0C0V0
B	451	HIS	-	expression tag	UNP P0C0V0
B	452	HIS	-	expression tag	UNP P0C0V0
B	453	HIS	-	expression tag	UNP P0C0V0
B	454	HIS	-	expression tag	UNP P0C0V0
B	455	HIS	-	expression tag	UNP P0C0V0
B	456	HIS	-	expression tag	UNP P0C0V0

- Molecule 2 is DIISOPROPYL PHOSPHONATE (three-letter code: DFP) (formula: C₆H₁₅O₃P).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	P		
2	A	1	10	6	3	1	0	0
2	B	1	10	6	3	1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	2	2	2	0	0
3	B	7	7	7	0	0

F240	ILE	ILE	SER	VAL
A241	LYS	LYS	SER	LEU
T242	ALA	SER	SER	ASP
P243	GLY	ILE	SER	ILE
S244	ASP	PHE	LYS	PRO
N245	VAL	ASN	ASN	SER
M246	ILE	GLY	GLY	VAL
	THR	ILE	ILE	LEU
	THR	THR	GLY	ALA
L250	SER	LEU	GLY	LEU
	LEU	GLY	ALA	LEU
Q253	ASN	ALA	ASN	ASN
M254	GLY	GLY	GLY	ILE
	LYS	MET	MET	GLN
	PRO	SER	SER	GLN
Y257	ILE	ASN	ARG	ARG
G258	SER	LYS	GLY	GLY
Q259	SER	LYS	ASP	SER
V260	ASP	GLY	SER	SER
K261	PHE	LYS	THR	THR
R262	ALA	ASP	GLN	ILE
G263	ALA	LEU	GLY	TYR
E264	LEU	ARG	VAL	LEU
	GLY	ILE	VAL	LEU
	ILE	GLN	VAL	LEU
	ILE	GLN	ASN	MET
	MET	VAL	ASN	GLN
	GLY	GLY	ASN	ARG
	THR	THR	VAL	SER
	GLU	MET	LYS	HIS
	LEU	PRO	THR	HIS
	ASN	VAL	GLY	HIS
	SER	GLY	THR	HIS
	GLU	GLY	THR	HIS
	LEU	SER	PRO	HIS
	LEU	LYS	ALA	HIS
	ALA	ALA	ALA	
	LYS	LEU	GLN	
	ALA	THR	ILE	
	LYS	THR	GLY	
	ALA	GLY	LEU	
	MET	LEU	LEU	
	LYS	ARG	LYS	
	VAL	ASP	GLY	
	ARG	ALA	ASP	
	GLY	GLN	VAL	
	ARC	LYS	ILE	
	GLY	GLN	ILE	
	ALA	VAL	ILE	
	PHE	ASN	GLY	
	VAL	VAL	ALA	
	SER	ASN	ASN	
	SER	LEU	GLN	
	VAL	GLY	GLY	
	LEU	LEU	ALA	
	PRO	LEU	VAL	
	ASN	GLN	VAL	
	ASN	GLN	LYS	
	SER	SER	ILE	
	ALA	GLN	ALA	
	LYS	ASN	GLY	
	ALA	GLN	LEU	
	GLY	VAL	ARG	
	GLY	ASP	LYS	

4 Data and refinement statistics

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α , β , γ	121.45Å 121.45Å 226.65Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.00 24.94 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.1 (20.00-3.00) 99.2 (24.94-3.00)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.46 (at 2.99Å)	Xtrriage
Refinement program	CNS	Depositor
R, R_{free}	0.281 , 0.309 0.276 , 0.304	Depositor DCC
R_{free} test set	942 reflections (4.60%)	wwPDB-VP
Wilson B-factor (Å ²)	85.1	Xtrriage
Anisotropy	0.718	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 99.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	2922	wwPDB-VP
Average B, all atoms (Å ²)	97.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.32% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: DFP

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	0.37	0/1422	0.65	0/1920
1	B	0.38	0/1497	0.63	0/2021
All	All	0.38	0/2919	0.64	0/3941

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	1409	0	1436	79	0
1	B	1484	0	1515	65	0
2	A	10	0	14	1	0
2	B	10	0	14	1	0
3	A	2	0	0	0	0
3	B	7	0	0	0	0
All	All	2922	0	2979	142	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 24.

The worst 5 of 142 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:B:104:ASN:HD21	1:B:130:LYS:HE3	1.23	1.04
1:A:104:ASN:HD21	1:A:130:LYS:HE3	1.27	0.99
1:B:253:GLN:HG2	1:B:260:VAL:HA	1.52	0.91
1:B:171:PHE:HB2	1:B:173:LEU:HD12	1.52	0.91
1:A:198:PHE:HB3	1:A:243:PRO:HD3	1.51	0.90

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	183/456 (40%)	169 (92%)	11 (6%)	3 (2%)	9	40
1	B	194/456 (42%)	179 (92%)	14 (7%)	1 (0%)	29	68
All	All	377/912 (41%)	348 (92%)	25 (7%)	4 (1%)	14	50

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	THR
1	A	84	PHE
1	A	161	GLY
1	B	161	GLY

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	155/365 (42%)	140 (90%)	15 (10%)	8	31
1	B	163/365 (45%)	149 (91%)	14 (9%)	10	37
All	All	318/730 (44%)	289 (91%)	29 (9%)	9	34

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	246	MET
1	B	217	LEU
1	B	80	GLN
1	B	175	GLU
1	B	18	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	206	ASN
1	B	249	ASN
1	B	259	GLN
1	B	245	ASN
1	B	80	GLN

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](#)

2 ligands are modelled in this entry.

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	$\# Z > 2$
2	DFP	B	500	1	6,9,9	1.18	0	6,11,11	0.41	0
2	DFP	A	500	1	6,9,9	1.28	0	6,11,11	0.42	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	DFP	B	500	1	-	0/4/8/8	-
2	DFP	A	500	1	-	0/4/8/8	-

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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	500	DFP	1	0
2	A	500	DFP	1	0

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	191/456 (41%)	0.25	16 (8%) 11 3	63, 91, 156, 197	0
1	B	202/456 (44%)	0.12	16 (7%) 12 4	59, 85, 151, 181	0
All	All	393/912 (43%)	0.18	32 (8%) 12 3	59, 88, 156, 197	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	11	GLN	4.8
1	A	263	GLY	4.0
1	A	197	ASN	3.9
1	A	206	ASN	3.9
1	A	228	ILE	3.8

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	DFP	A	500	10/10	0.90	0.26	172,173,175,175	0
2	DFP	B	500	10/10	0.93	0.21	113,118,120,121	0

6.5 Other polymers [i](#)

There are no such residues in this entry.