



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 04:35 am BST

PDB ID : 2P3W  
Title : Crystal Structure of the HtrA3 PDZ Domain Bound to a Phage-Derived Ligand (FGRWV)  
Authors : Appleton, B.A.; Wiesmann, C.  
Deposited on : 2007-03-09  
Resolution : 1.70 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](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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

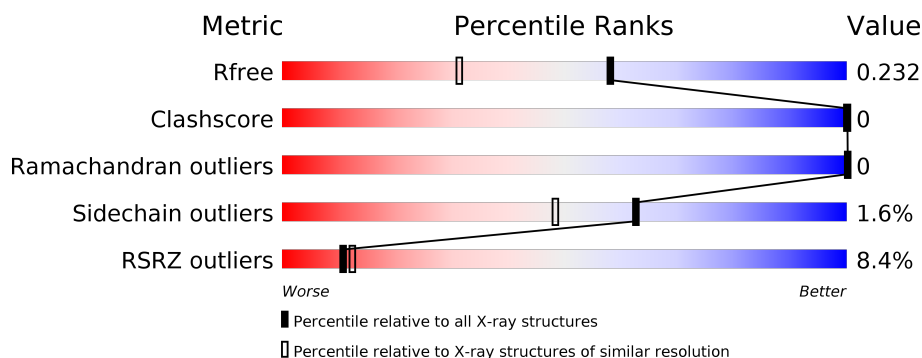
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 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  $\leq 5\%$ . 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	112	<div> <div>10%</div> <div> <div></div> <div>98%</div> <div>.</div> </div> </div>
1	B	112	<div> <div>6%</div> <div> <div></div> <div>90%</div> <div>7%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1854 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 Probable serine protease HTRA3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	110	Total	C	N	O	S	0	1	0
			842	528	148	163	3			
1	B	104	Total	C	N	O	S	0	2	0
			805	506	142	154	3			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	350	GLY	-	EXPRESSION TAG	UNP P83110
A	351	SER	-	EXPRESSION TAG	UNP P83110
A	352	HIS	-	EXPRESSION TAG	UNP P83110
A	353	MET	-	EXPRESSION TAG	UNP P83110
A	454	GLY	-	LINKER	UNP P83110
A	455	GLY	-	LINKER	UNP P83110
A	456	GLY	-	LINKER	UNP P83110
A	457	PHE	-	SEE REMARK 999	UNP P83110
A	458	GLY	-	SEE REMARK 999	UNP P83110
A	459	ARG	-	SEE REMARK 999	UNP P83110
A	460	TRP	-	SEE REMARK 999	UNP P83110
A	461	VAL	-	SEE REMARK 999	UNP P83110
B	350	GLY	-	EXPRESSION TAG	UNP P83110
B	351	SER	-	EXPRESSION TAG	UNP P83110
B	352	HIS	-	EXPRESSION TAG	UNP P83110
B	353	MET	-	EXPRESSION TAG	UNP P83110
B	454	GLY	-	LINKER	UNP P83110
B	455	GLY	-	LINKER	UNP P83110
B	456	GLY	-	LINKER	UNP P83110
B	457	PHE	-	SEE REMARK 999	UNP P83110
B	458	GLY	-	SEE REMARK 999	UNP P83110
B	459	ARG	-	SEE REMARK 999	UNP P83110
B	460	TRP	-	SEE REMARK 999	UNP P83110
B	461	VAL	-	SEE REMARK 999	UNP P83110

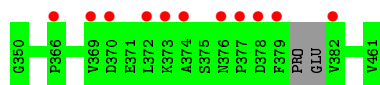
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	109	Total 109	O 109	0	0
2	B	98	Total 98	O 98	0	0

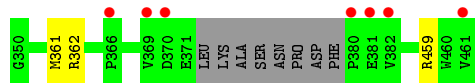
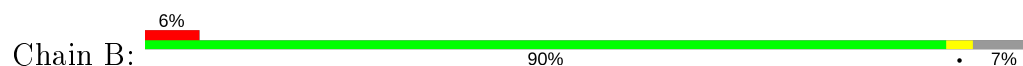
### 3 Residue-property plots [i](#)

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 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 ( $\text{RSRZ} > 2$ ). 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: Probable serine protease HTA3



- Molecule 1: Probable serine protease HTA3



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.00Å 73.00Å 80.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.70 35.10 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (30.00-1.70) 99.9 (35.10-1.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.94 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.182 , 0.220 0.195 , 0.232	Depositor DCC
$R_{free}$ test set	1224 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.8	Xtriage
Anisotropy	0.253	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 44.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	1854	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 55.29 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.2310e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.59	0/858	0.71	0/1158
1	B	0.58	0/824	0.73	0/1112
All	All	0.59	0/1682	0.72	0/2270

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	842	0	848	0	0
1	B	805	0	815	0	0
2	A	109	0	0	0	0
2	B	98	0	0	0	0
All	All	1854	0	1663	0	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 0.

There are no clashes within the asymmetric unit.

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	107/112 (96%)	106 (99%)	1 (1%)	0	100	100
1	B	102/112 (91%)	102 (100%)	0	0	100	100
All	All	209/224 (93%)	208 (100%)	1 (0%)	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	95/96 (99%)	95 (100%)	0	100	100
1	B	91/96 (95%)	88 (97%)	3 (3%)	38	19
All	All	186/192 (97%)	183 (98%)	3 (2%)	62	48

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	361	MET
1	B	362	ARG
1	B	459	ARG

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



### 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 carbohydrates 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	110/112 (98%)	0.45	11 (10%) 7 8	15, 20, 42, 52	0
1	B	104/112 (92%)	0.38	7 (6%) 17 20	14, 20, 36, 53	0
All	All	214/224 (95%)	0.42	18 (8%) 11 12	14, 20, 41, 53	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	382	VAL	6.0
1	A	378	ASP	5.6
1	A	374	ALA	5.5
1	A	373	LYS	5.2
1	A	377	PRO	5.0

### 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 carbohydrates 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.