



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 07:25 am GMT

PDB ID : 1KY9  
Title : Crystal Structure of DegP (HtrA)  
Authors : Krojer, T.; Garrido-Franco, M.; Huber, R.; Ehrmann, M.; Clausen, T.  
Deposited on : 2002-02-04  
Resolution : 2.80 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

<http://wwpdb.org/validation/2016/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)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

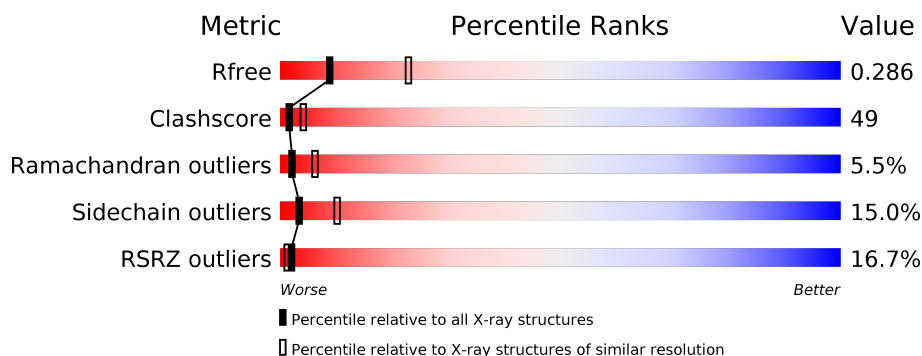
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
$R_{free}$	100719	2583 (2.80-2.80)
Clashscore	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)
RSRZ outliers	101464	2610 (2.80-2.80)

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. 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	448	<div> <div>12%</div> <div> <div>32%</div> <div>27%</div> <div>9%</div> <div>31%</div> </div> </div>
1	B	448	<div> <div>13%</div> <div> <div>35%</div> <div>44%</div> <div>9%</div> <div>12%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5366 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
1	A	311	Total	C	N	O	Se	0	0	0
			2282	1426	403	441	12			
1	B	396	Total	C	N	O	Se	0	0	0
			2918	1818	519	568	13			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	12	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	18	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	23	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	42	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	85	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	127	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	152	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	210	ALA	SER	ENGINEERED	UNP P0C0V0
A	246	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	254	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	268	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	280	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	331	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	376	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
A	447	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	12	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	18	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	23	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	42	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	85	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	127	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	152	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	210	ALA	SER	ENGINEERED	UNP P0C0V0
B	246	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	254	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	268	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	280	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	331	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	376	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0
B	447	MSE	MET	MODIFIED RESIDUE	UNP P0C0V0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	77	Total O 77 77	0	0
2	B	89	Total O 89 89	0	0



E420	L421	R422	K423	V424	L425	D426	S427	K428	P429	S430	V431	L432	A433	L434	N435	I436	Q437	R438	G439	D440	S441	T442	I443	Y444	L445	L446	MSE	GLN									
Q359	N360	Q361	V362	D363	S364	S365	S366	I367	F368	N369	GLY	ILE	GLU	GLY	ALA	E375	M376	S377	N378	K379	G380	K381	D382	Q383	G384	V385	V386	V387	N388	N389	V390	K391	T392	G393	T394	P395	A396
S297	S298	A299	A300	K301	A302	G303	T304	G307	D308	V309	I310	T311	S312	I313	K316	P317	I318	S319	S320	F321	A322	A323	L324	R325	A326	Q327	V328	T330	M331	V332	G333	G334	S335	K336	L337	T338	
G234	N235	I236	G237	I238	G239	F240	P243	S244	N245	M246	V247	K248	N249	L250	T251	M254	V255	E256	Y257	G258	Q259	V260	K261	R262	L265	G266	L267	M268	G269	T270	E271	L272	N273	S274	E275	L276	
A277	K278	A279	M280	K281	V282	D283	A284	Q285	K286	G287	A288	F289	V290	S291	Q292	V293	L294	P295	N296																		

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	121.37Å 121.37Å 233.67Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.80 19.93 – 2.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.80) 99.6 (19.93-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.48 (at 2.79Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.218 , 0.275 0.228 , 0.286	Depositor DCC
$R_{free}$ test set	1273 reflections (4.96%)	DCC
Wilson B-factor (Å <sup>2</sup> )	69.7	Xtriage
Anisotropy	0.398	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 83.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5366	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.40% of the height of the origin peak. No significant pseudotranslation is detected.*

<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	0.61	0/2295	0.88	7/3078 (0.2%)
1	B	0.56	0/2932	0.90	4/3934 (0.1%)
All	All	0.58	0/5227	0.89	11/7012 (0.2%)

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	230	ALA	C-N-CD	-20.86	74.72	120.60
1	B	230	ALA	C-N-CA	13.68	179.44	122.00
1	A	230	ALA	C-N-CD	-9.74	99.18	120.60
1	A	230	ALA	C-N-CA	7.04	151.57	122.00
1	B	87	LEU	CA-CB-CG	6.67	130.65	115.30

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	2282	0	2340	225	0
1	B	2918	0	3008	290	0
2	A	77	0	0	18	0
2	B	89	0	0	15	0
All	All	5366	0	5348	515	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 49.

The worst 5 of 515 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:266:GLY:O	1:A:267:ILE:HG13	1.38	1.21
1:B:267:ILE:HD11	1:B:321:PHE:HE1	1.15	1.08
1:A:309:VAL:HG23	1:A:342:LEU:HB3	1.33	1.06
1:B:246:MSE:HG2	2:B:476:HOH:O	1.56	1.05
1:A:265:LEU:HG	1:A:265:LEU:O	1.28	1.04

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	305/448 (68%)	259 (85%)	29 (10%)	17 (6%)	2	6
1	B	388/448 (87%)	318 (82%)	49 (13%)	21 (5%)	2	6
All	All	693/896 (77%)	577 (83%)	78 (11%)	38 (6%)	2	6

5 of 38 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	197	ASN
1	A	267	ILE
1	A	268	MSE
1	A	271	GLU
1	A	281	LYS

### 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	246/342 (72%)	207 (84%)	39 (16%)	3	9
1	B	319/342 (93%)	273 (86%)	46 (14%)	4	11
All	All	565/684 (83%)	480 (85%)	85 (15%)	3	10

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

Mol	Chain	Res	Type
1	A	343	ARG
1	B	93	ILE
1	B	354	LEU
1	B	12	MSE
1	B	49	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	104	ASN
1	B	146	ASN
1	B	388	ASN
1	B	140	GLN
1	B	169	ASN

### 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	299/448 (66%)	0.74	54 (18%) <b>1</b> <b>1</b>	35, 68, 175, 186	0
1	B	383/448 (85%)	0.65	60 (15%) <b>2</b> <b>1</b>	37, 102, 163, 169	0
All	All	682/896 (76%)	0.69	114 (16%) <b>2</b> <b>1</b>	35, 83, 167, 186	0

The worst 5 of 114 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	439	GLY	10.9
1	A	345	GLY	8.3
1	A	274	SER	8.1
1	A	271	GLU	7.9
1	A	276	LEU	7.8

### 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 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.