



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 11:53 PM GMT

PDB ID : 1YZE  
Title : Crystal structure of the N-terminal domain of USP7/HAUSP.  
Authors : Saridakis, V.; Sheng, Y.; Sarkari, F.; Holowaty, M.N.; Shire, K.; Nguyen, T.; Zhang, R.G.; Liao, J.; Lee, W.; Edwards, A.M.; Arrowsmith, C.H.; Frappier, L.  
Deposited on : 2005-02-28  
Resolution : 2.00 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](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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

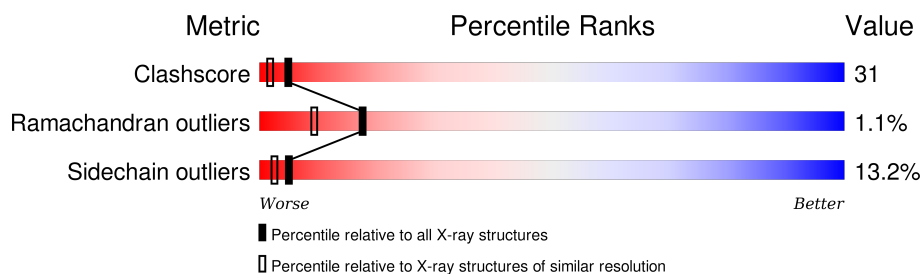
# 1 Overall quality at a glance

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	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	155	<div> <div>27%</div> <div>34%</div> <div>•</div> <div>35%</div> </div>
1	B	155	<div> <div>26%</div> <div>32%</div> <div>6%</div> <div>36%</div> </div>
1	C	155	<div> <div>30%</div> <div>29%</div> <div>•</div> <div>37%</div> </div>

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 2537 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 Ubiquitin carboxyl-terminal hydrolase 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	101	Total	C	N	O	S	0	0	0
			840	552	144	138	6			
1	B	99	Total	C	N	O	S	0	0	0
			813	533	138	136	6			
1	C	98	Total	C	N	O	S	0	0	0
			794	519	132	137	6			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	51	GLY	-	cloning artifact	UNP Q93009
A	52	SER	-	cloning artifact	UNP Q93009
A	53	HIS	-	cloning artifact	UNP Q93009
B	51	GLY	-	cloning artifact	UNP Q93009
B	52	SER	-	cloning artifact	UNP Q93009
B	53	HIS	-	cloning artifact	UNP Q93009
C	51	GLY	-	cloning artifact	UNP Q93009
C	52	SER	-	cloning artifact	UNP Q93009
C	53	HIS	-	cloning artifact	UNP Q93009

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	35	Total	O	0	0
			35	35		
2	B	25	Total	O	0	0
			25	25		
2	C	30	Total	O	0	0
			30	30		



Note EDS was not executed.

- Chain A:
- 
- 27% 34% 35%
- GLY SER HIS THR ALA GLU ASP MET GLU ASP THR W65
- F17 F18 F19 C120 C121 M122 A123 E124 T128 S129 W130 Q135 A136 V137 L138 K139 I140 I141 A142 T143 V144 A145 A146 A147 A148 A149 A150 A151 A152 A153 F159 H160 K161 E162 M163 D164 F167 S168 M169 F170 M171 A172 W173 SER GLU VAL THR PRO PRO GLU LYS GLY PHE ILE ASP W1145

- Chain B:
- 
- |      |      |      |      |      |      |      |      |      |      |     |     |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |     |     |     |     |     |     |     |     |      |
|------|------|------|------|------|------|------|------|------|------|-----|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-----|-----|-----|-----|-----|-----|-----|-----|------|
| GLY  | SER  | HIS  | THR  | ALA  | GLU  | ASP  | MET  | GLU  | ASP  | THR | SER | TRP  | ARG  | S67  |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |     |     |     |     |     |     |     |     |      |
| LYS  | SER  | V115 | G116 | F117 | F118 | L119 | Q120 | E124 | S125 | ASP | THR | T129 | W130 | S131 |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |     |     |     |     |     |     |     |     |      |
| Q135 | A136 | K139 | I140 | I141 | I142 | TYR  | ARG  | ASP  | ASP  | ASP | GLU | LYS  | SER  | F150 | S151 | R152 | R153 | I154 | S155 | F158 | F159 | H160 | K161 | E162 | M163 | D164 | M171 | A172 | W173 | S174 | E175 | VAL | THR | ASP | PRO | GLU | LYS | PRO | GLN | P176 |

- Chain C:
- GLY SER HIS THR ALA GLU GLU ASP MET GLU ASP ASP THR TRP ARG S67 F73 T74 VAL GLU ARG ARG PHE SER ARG LEU SER GLU SER Y85 L86 S87 R93 I94 K98 I99 M100 Y101 M102 PRU ARG ARG PHE TVR PRO ASP ARG PRO HIS GLN LYS SER V115 G116 F117 F118 L119 Q120 C121 M122 E123 E124 D125 D126 S127 T128 S129 W130 S131 A136 V137 I138 K139 T140 I141 M142 TYR ARG ASP ASP GLU LYS SER F150 S151 R152 R153 I154 S155 L156 L157 K161 E162 M163 D164 W165 G166 F170 M171 A172 E176 V176 THR ASP PRO GLU LYS GLY PHE ILE ASP ASP ASP A158 T190 E191 E192 V193 F194 P200 H201 G202 V203 A204 TRP

## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.61Å 102.61Å 45.20Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	26.96 – 2.00	Depositor
% Data completeness (in resolution range)	98.2 (26.96-2.00)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.325 , 0.367	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2537	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

## 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	0.52	0/867	0.70	0/1170
1	B	0.52	0/837	0.67	0/1128
1	C	0.49	0/817	0.74	0/1103
All	All	0.51	0/2521	0.70	0/3401

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	840	0	801	51	0
1	B	813	0	770	60	0
1	C	794	0	752	38	0
2	A	35	0	0	6	0
2	B	25	0	0	3	0
2	C	30	0	0	4	0
All	All	2537	0	2323	149	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 31.

The worst 5 of 149 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:C:172:ALA:HB3	1:C:175:GLU:HB2	1.48	0.94
1:C:100:MET:HG2	1:C:118:PHE:HB2	1.54	0.89
1:B:151:SER:O	1:B:152:ARG:HG3	1.79	0.82
1:A:139:LYS:HD3	1:A:141:ILE:HD13	1.61	0.80
1:B:72:GLN:NE2	1:B:192:GLU:HG3	1.97	0.80

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	91/155 (59%)	82 (90%)	8 (9%)	1 (1%)	17	9
1	B	87/155 (56%)	81 (93%)	5 (6%)	1 (1%)	17	9
1	C	88/155 (57%)	81 (92%)	6 (7%)	1 (1%)	17	9
All	All	266/465 (57%)	244 (92%)	19 (7%)	3 (1%)	17	9

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	125	SER
1	B	161	LYS
1	A	202	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	91/141 (64%)	79 (87%)	12 (13%)	5	2
1	B	88/141 (62%)	77 (88%)	11 (12%)	6	3
1	C	87/141 (62%)	75 (86%)	12 (14%)	4	2
All	All	266/423 (63%)	231 (87%)	35 (13%)	5	2

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	85	VAL
1	B	135	GLN
1	C	164	ASP
1	B	102	MET
1	B	117	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	160	HIS
1	B	163	ASN
1	C	94	ASN
1	B	72	GLN
1	C	120	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 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](#)

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.