



# Full wwPDB X-ray Structure Validation Report ⓘ

May 11, 2021 – 04:01 pm BST

PDB ID : 6Z0E  
Title : HtrA1 inactive protease domain S328A with CARASIL mutation R274Q  
Authors : Vetter, I.R.; Stege, P.; Ingendahl, L.; Ehrmann, M.  
Deposited on : 2020-05-08  
Resolution : 2.60 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  
Xtriage (Phenix) : 1.13  
EDS : 2.18  
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) : 2.18

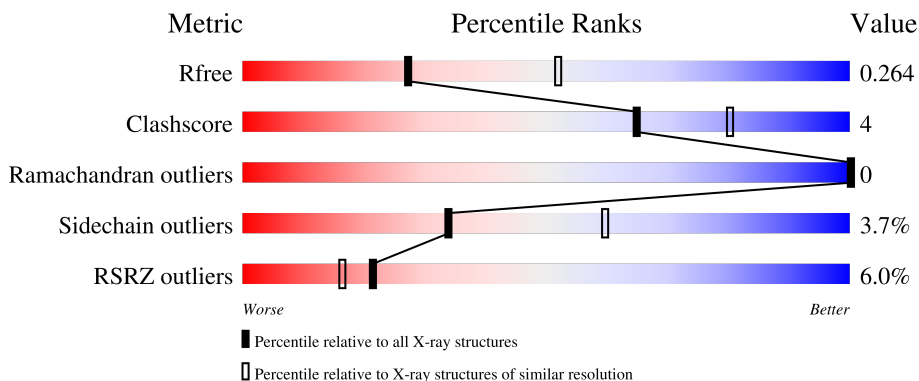
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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  $\geq 3$ , 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	236	 3% 72% 10% 17%
1	B	236	 6% 68% 12% 20%

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 2961 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 Serine protease HTRA1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	196	1497	957	254	286	0	0	0
1	B	188	1429	912	242	275	0	0	0

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	140	MET	-	initiating methionine	UNP Q92743
A	141	GLY	-	expression tag	UNP Q92743
A	142	SER	-	expression tag	UNP Q92743
A	143	SER	-	expression tag	UNP Q92743
A	144	HIS	-	expression tag	UNP Q92743
A	145	HIS	-	expression tag	UNP Q92743
A	146	HIS	-	expression tag	UNP Q92743
A	147	HIS	-	expression tag	UNP Q92743
A	148	HIS	-	expression tag	UNP Q92743
A	149	HIS	-	expression tag	UNP Q92743
A	150	SER	-	expression tag	UNP Q92743
A	151	SER	-	expression tag	UNP Q92743
A	152	GLY	-	expression tag	UNP Q92743
A	153	LEU	-	expression tag	UNP Q92743
A	154	VAL	-	expression tag	UNP Q92743
A	155	PRO	-	expression tag	UNP Q92743
A	156	ARG	-	expression tag	UNP Q92743
A	157	GLY	-	expression tag	UNP Q92743
A	158	SER	-	expression tag	UNP Q92743
A	159	HIS	-	expression tag	UNP Q92743
A	160	MET	-	expression tag	UNP Q92743
A	274	GLN	ARG	engineered mutation	UNP Q92743
A	328	ALA	SER	engineered mutation	UNP Q92743
B	140	MET	-	initiating methionine	UNP Q92743
B	141	GLY	-	expression tag	UNP Q92743

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	142	SER	-	expression tag	UNP Q92743
B	143	SER	-	expression tag	UNP Q92743
B	144	HIS	-	expression tag	UNP Q92743
B	145	HIS	-	expression tag	UNP Q92743
B	146	HIS	-	expression tag	UNP Q92743
B	147	HIS	-	expression tag	UNP Q92743
B	148	HIS	-	expression tag	UNP Q92743
B	149	HIS	-	expression tag	UNP Q92743
B	150	SER	-	expression tag	UNP Q92743
B	151	SER	-	expression tag	UNP Q92743
B	152	GLY	-	expression tag	UNP Q92743
B	153	LEU	-	expression tag	UNP Q92743
B	154	VAL	-	expression tag	UNP Q92743
B	155	PRO	-	expression tag	UNP Q92743
B	156	ARG	-	expression tag	UNP Q92743
B	157	GLY	-	expression tag	UNP Q92743
B	158	SER	-	expression tag	UNP Q92743
B	159	HIS	-	expression tag	UNP Q92743
B	160	MET	-	expression tag	UNP Q92743
B	274	GLN	ARG	engineered mutation	UNP Q92743
B	328	ALA	SER	engineered mutation	UNP Q92743

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Ca 1 1	0	0
2	B	1	Total Ca 1 1	0	0

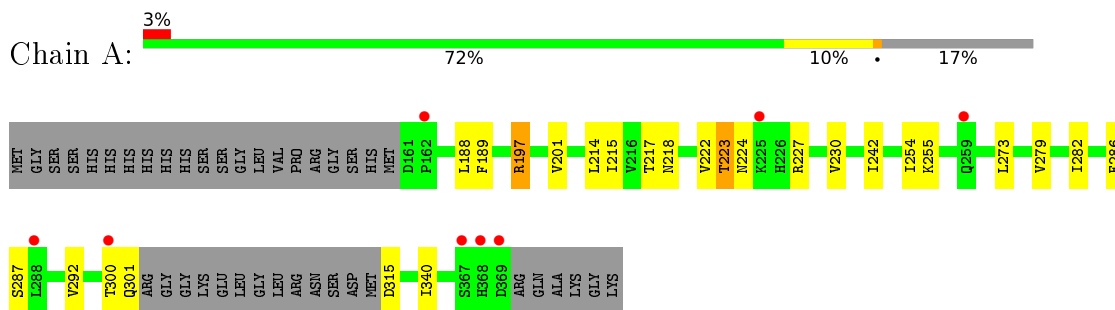
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	20	Total O 20 20	0	0
3	B	13	Total O 13 13	0	0

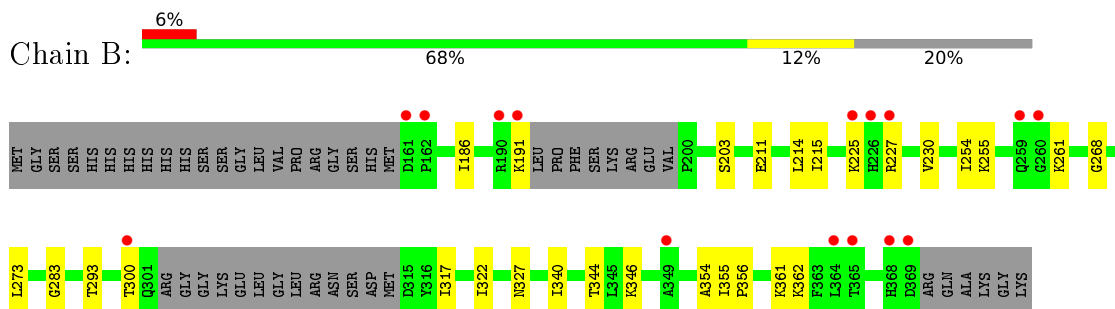
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 ( $RSRZ > 2$ ). 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: Serine protease HTRA1



- Molecule 1: Serine protease HTRA1



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	109.86Å 109.86Å 114.40Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.02 – 2.60 49.02 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.02-2.60) 100.0 (49.02-2.60)	Depositor EDS
$R_{merge}$	0.33	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.04 (at 2.61Å)	Xtrriage
Refinement program	PHENIX 1.18rc2_3794	Depositor
R, $R_{free}$	0.238 , 0.265 0.237 , 0.264	Depositor DCC
$R_{free}$ test set	790 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	54.9	Xtrriage
Anisotropy	0.037	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 35.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.026 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	2961	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.71% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.37	0/1522	0.58	0/2064
1	B	0.39	0/1451	0.66	2/1966 (0.1%)
All	All	0.38	0/2973	0.62	2/4030 (0.0%)

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	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	227	ARG	NE-CZ-NH2	7.85	124.23	120.30
1	B	227	ARG	NE-CZ-NH1	-5.17	117.72	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	223	THR	Peptide

### 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	1497	0	1536	11	0
1	B	1429	0	1463	13	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	20	0	0	0	0
3	B	13	0	0	1	0
All	All	2961	0	2999	24	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 4.

All (24) 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:273:LEU:HD11	1:A:340:ILE:HD13	1.88	0.55
1:A:282:ILE:HG13	1:A:292:VAL:HG22	1.92	0.51
1:B:215:ILE:HD13	1:B:230:VAL:HG11	1.92	0.51
1:B:215:ILE:HB	1:B:254:ILE:HG13	1.93	0.50
1:B:273:LEU:HD11	1:B:340:ILE:HD13	1.94	0.49
1:B:317:ILE:HB	1:B:354:ALA:HB3	1.94	0.49
1:A:189:PHE:HB2	1:A:227:ARG:HB3	1.94	0.49
1:B:268:GLY:O	1:B:361:LYS:NZ	2.44	0.48
1:A:315:ASP:OD2	1:A:315:ASP:N	2.49	0.46
1:A:217:THR:OG1	1:A:218:ASN:N	2.50	0.45
1:B:283:GLY:HA3	1:B:327:ASN:HA	1.99	0.44
1:A:242:ILE:HA	1:A:254:ILE:HG22	1.98	0.44
1:B:344:THR:HG22	1:B:346:LYS:H	1.83	0.44
1:B:293:THR:HG21	1:B:322:ILE:HB	1.99	0.43
1:A:215:ILE:HD13	1:A:230:VAL:HG11	2.00	0.43
1:B:214:LEU:HD23	1:B:255:LYS:HB2	2.02	0.42
1:B:186:ILE:HG12	1:B:230:VAL:HG22	2.01	0.41
1:B:211:GLU:HB2	1:B:261:LYS:HD2	2.02	0.41
1:A:197:ARG:HE	1:A:197:ARG:HB2	1.67	0.41
1:A:223:THR:OG1	1:A:224:ASN:N	2.53	0.41
1:A:286:PHE:HB3	1:A:287:SER:H	1.71	0.41
1:B:346:LYS:NZ	3:B:504:HOH:O	2.54	0.41
1:A:214:LEU:HD23	1:A:255:LYS:HB2	2.02	0.40
1:B:355:ILE:HA	1:B:356:PRO:HD3	1.91	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	192/236 (81%)	184 (96%)	8 (4%)	0	100	100
1	B	182/236 (77%)	173 (95%)	9 (5%)	0	100	100
All	All	374/472 (79%)	357 (96%)	17 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	167/199 (84%)	160 (96%)	7 (4%)	30	55
1	B	159/199 (80%)	154 (97%)	5 (3%)	40	66
All	All	326/398 (82%)	314 (96%)	12 (4%)	34	60

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	188	LEU
1	A	197	ARG
1	A	201	VAL
1	A	222	VAL
1	A	279	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	300	THR
1	A	301	GLN
1	B	191	LYS
1	B	203	SER
1	B	225	LYS
1	B	300	THR
1	B	362	LYS

Sometimes 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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

### 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

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	196/236 (83%)	0.45	8 (4%) 37 30	29, 51, 86, 107	0
1	B	188/236 (79%)	0.57	15 (7%) 12 9	37, 61, 94, 108	0
All	All	384/472 (81%)	0.51	23 (5%) 21 16	29, 57, 90, 108	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	368	HIS	7.2
1	A	369	ASP	4.9
1	B	225	LYS	4.9
1	B	162	PRO	4.5
1	A	225	LYS	4.5
1	B	369	ASP	3.8
1	B	368	HIS	3.5
1	A	162	PRO	3.4
1	B	365	THR	3.2
1	B	191	LYS	3.1
1	B	227	ARG	3.1
1	B	300	THR	3.1
1	B	226	HIS	3.0
1	A	300	THR	3.0
1	A	288	LEU	2.9
1	B	161	ASP	2.7
1	A	367	SER	2.5
1	B	259	GLN	2.4
1	B	260	GLY	2.3
1	A	259	GLN	2.2
1	B	364	LEU	2.1
1	B	349	ALA	2.0
1	B	190	ARG	2.0

## 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 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	CA	B	401	1/1	0.69	0.15	48,48,48,48	1
2	CA	A	401	1/1	0.98	0.17	42,42,42,42	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.