



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 21, 2017 – 06:27 AM EST

PDB ID : 3ER5  
Title : THE ACTIVE SITE OF ASPARTIC PROTEINASES  
Authors : Bailey, D.; Veerapandian, B.; Cooper, J.; Szelke, M.; Blundell, T.L.  
Deposited on : unknown  
Resolution : 1.80 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

<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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

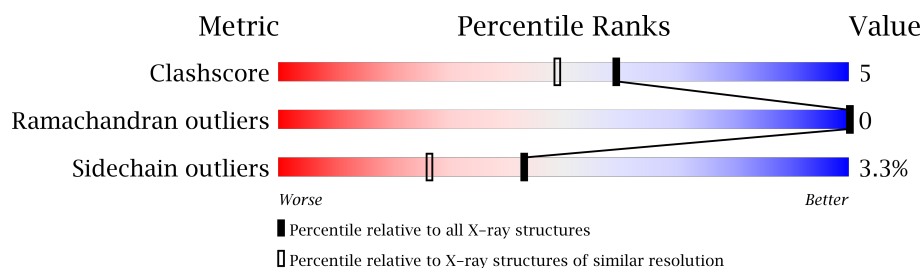
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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))
Clashscore	112137	5742 (1.80-1.80)
Ramachandran outliers	110173	5676 (1.80-1.80)
Sidechain outliers	110143	5675 (1.80-1.80)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	E	330	
2	I	10	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2703 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 ENDOTHAPEPSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	330	Total	C	N	O	S	0	0	0
			2389	1514	366	507	2			

- Molecule 2 is a protein called H-189.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	I	10	Total	C	N	O	0	0	0
			91	62	17	12			

- Molecule 3 is water.

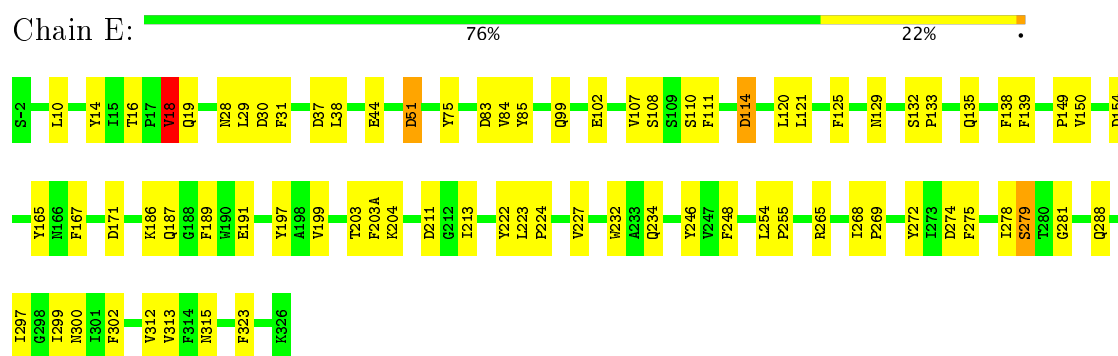
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	217	Total	O	0	0
			217	217		
3	I	6	Total	O	0	0
			6	6		

### 3 Residue-property plots

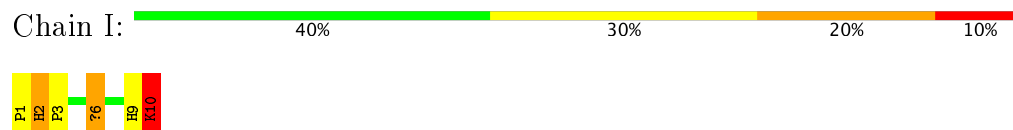
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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 ( $RSRZ > 2$ ). 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.

Note EDS was not executed.

#### • Molecule 1: ENDOTHIAPEPSIN



#### • Molecule 2: H-189



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	43.10 Å 75.40 Å 42.80 Å 90.00° 97.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.80	Depositor
% Data completeness (in resolution range)	(Not available) (10.00-1.80)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, $R_{free}$	0.150 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2703	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	0.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

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

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	E	1.02	7/2445 (0.3%)	1.64	45/3345 (1.3%)
2	I	1.66	2/84 (2.4%)	2.29	7/110 (6.4%)
All	All	1.05	9/2529 (0.4%)	1.66	52/3455 (1.5%)

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	E	0	2
2	I	0	3
All	All	0	5

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	269	PRO	N-CD	8.55	1.59	1.47
2	I	3	PRO	N-CD	8.48	1.59	1.47
1	E	274	ASP	C-N	8.46	1.53	1.34
1	E	18	VAL	C-N	7.15	1.50	1.34
1	E	268	ILE	C-N	6.52	1.46	1.34
1	E	232	TRP	NE1-CE2	-5.92	1.29	1.37
1	E	149	PRO	C-N	-5.71	1.21	1.34
2	I	10	LYS	C-OXT	5.25	1.33	1.23
1	E	19	GLN	CD-OE1	5.21	1.35	1.24

All (52) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	114	ASP	CB-CG-OD2	-13.76	105.92	118.30
2	I	2	HIS	C-N-CD	13.18	156.07	128.40
1	E	30	ASP	CB-CG-OD1	11.45	128.60	118.30
1	E	323	PHE	CB-CG-CD2	-9.35	114.26	120.80
1	E	114	ASP	OD1-CG-OD2	9.02	140.45	123.30
1	E	83	ASP	CB-CG-OD2	9.02	126.42	118.30
1	E	274	ASP	C-N-CA	-8.95	99.33	121.70
1	E	268	ILE	C-N-CD	8.33	145.89	128.40
1	E	246	TYR	CB-CG-CD1	-8.15	116.11	121.00
1	E	75	TYR	CB-CG-CD1	-7.83	116.30	121.00
1	E	44	GLU	OE1-CD-OE2	-7.82	113.91	123.30
1	E	246	TYR	CG-CD2-CE2	-7.68	115.16	121.30
1	E	191	GLU	OE1-CD-OE2	-7.33	114.50	123.30
1	E	171	ASP	CB-CG-OD2	7.07	124.66	118.30
1	E	272	TYR	CB-CG-CD2	-6.86	116.88	121.00
2	I	9	HIS	CA-CB-CG	-6.78	102.07	113.60
2	I	3	PRO	CA-N-CD	-6.62	102.24	111.50
1	E	149	PRO	O-C-N	-6.55	112.22	122.70
1	E	14	TYR	CB-CG-CD2	-6.51	117.10	121.00
1	E	18	VAL	CA-C-N	-6.47	102.96	117.20
1	E	150	VAL	CA-CB-CG1	6.43	120.55	110.90
1	E	121	LEU	CB-CG-CD2	6.36	121.82	111.00
1	E	189	PHE	CB-CG-CD2	-6.17	116.48	120.80
2	I	3	PRO	N-CD-CG	-6.04	94.15	103.20
1	E	167	PHE	CG-CD2-CE2	-5.98	114.22	120.80
1	E	139	PHE	CB-CG-CD1	-5.91	116.66	120.80
1	E	149	PRO	CA-C-N	5.89	130.17	117.20
1	E	197	TYR	CB-CG-CD2	-5.88	117.47	121.00
1	E	125	PHE	CB-CG-CD2	5.85	124.90	120.80
1	E	312	VAL	CA-CB-CG2	5.79	119.58	110.90
1	E	154	ASP	CB-CG-OD2	5.68	123.42	118.30
1	E	85	TYR	CG-CD1-CE1	-5.64	116.79	121.30
1	E	125	PHE	CB-CG-CD1	-5.62	116.86	120.80
1	E	165	TYR	CB-CG-CD1	-5.62	117.63	121.00
1	E	84	VAL	CG1-CB-CG2	-5.62	101.91	110.90
1	E	302	PHE	CZ-CE2-CD2	-5.62	113.36	120.10
1	E	315	ASN	O-C-N	-5.60	113.68	123.20
1	E	114	ASP	CA-CB-CG	-5.59	101.09	113.40
1	E	227	VAL	O-C-N	-5.54	113.84	122.70
1	E	211	ASP	CB-CG-OD1	-5.52	113.33	118.30
1	E	85	TYR	CB-CG-CD1	-5.31	117.81	121.00
1	E	102	GLU	OE1-CD-OE2	5.30	129.66	123.30
1	E	323	PHE	CG-CD1-CE1	-5.29	114.98	120.80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	313	VAL	CA-CB-CG2	5.26	118.80	110.90
1	E	114	ASP	CB-CG-OD1	-5.23	113.59	118.30
2	I	2	HIS	C-N-CA	-5.21	100.10	122.00
2	I	9	HIS	O-C-N	-5.21	114.36	122.70
2	I	9	HIS	CA-C-N	5.20	128.64	117.20
1	E	51	ASP	CB-CG-OD1	5.19	122.97	118.30
1	E	203(A)	PHE	CG-CD1-CE1	-5.12	115.17	120.80
1	E	139	PHE	CG-CD2-CE2	-5.08	115.21	120.80
1	E	323	PHE	CD1-CG-CD2	5.00	124.80	118.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	18	VAL	Mainchain
1	E	265	ARG	Sidechain
2	I	1	PRO	Peptide
2	I	6	STA	Mainchain,Peptide

## 5.2 Too-close contacts ⓘ

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	E	2389	0	2279	25	1
2	I	91	0	93	2	1
3	E	217	0	0	1	0
3	I	6	0	0	0	0
All	All	2703	0	2372	26	1

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 5.

All (26) 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:E:107:VAL:CG1	1:E:111:PHE:HB2	2.11	0.81

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:107:VAL:HG13	1:E:111:PHE:HB2	1.72	0.70
1:E:129:ASN:ND2	1:E:135:GLN:H	1.93	0.66
1:E:10:LEU:HD22	1:E:275:PHE:O	1.99	0.63
1:E:224:PRO:HD3	1:E:300:ASN:ND2	2.18	0.58
1:E:213:ILE:HG23	1:E:299:ILE:HD11	1.86	0.56
1:E:248:PHE:CZ	1:E:254:LEU:HD11	2.42	0.55
1:E:223:LEU:HB3	1:E:224:PRO:HD2	1.91	0.52
1:E:120:LEU:HD21	2:I:6:STA:HD12	1.93	0.50
1:E:278:ILE:HG13	1:E:279:SER:N	2.27	0.50
1:E:255:PRO:HG3	3:E:428:HOH:O	2.12	0.49
1:E:199:VAL:HG11	1:E:234:GLN:HG3	1.95	0.48
1:E:38:LEU:C	1:E:38:LEU:HD23	2.37	0.45
2:I:2:HIS:C	2:I:2:HIS:CD2	2.90	0.45
1:E:203:THR:OG1	1:E:204:LYS:HE3	2.19	0.43
1:E:297:ILE:HG22	1:E:299:ILE:H	1.83	0.43
1:E:16:THR:HB	1:E:31:PHE:CE2	2.53	0.43
1:E:99:GLN:NE2	1:E:138:PHE:HA	2.34	0.43
1:E:10:LEU:HA	1:E:10:LEU:HD23	1.79	0.42
1:E:132:SER:OG	1:E:133:PRO:HA	2.20	0.41
1:E:213:ILE:HG23	1:E:299:ILE:CD1	2.49	0.41
1:E:18:VAL:HG11	1:E:29:LEU:HD12	2.02	0.41
1:E:107:VAL:HG11	1:E:111:PHE:HB2	1.97	0.40
1:E:222:TYR:HA	1:E:288:GLN:O	2.21	0.40
1:E:107:VAL:HG12	1:E:108:SER:O	2.22	0.40
1:E:28:ASN:HA	1:E:28:ASN:HD22	1.69	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:281:GLY:O	2:I:10:LYS:CG[2_546]	2.16	0.04

## 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	E	328/330 (99%)	324 (99%)	4 (1%)	0	100	100
2	I	7/10 (70%)	5 (71%)	2 (29%)	0	100	100
All	All	335/340 (98%)	329 (98%)	6 (2%)	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	E	263/263 (100%)	255 (97%)	8 (3%)	46	30
2	I	9/9 (100%)	8 (89%)	1 (11%)	7	1
All	All	272/272 (100%)	263 (97%)	9 (3%)	43	26

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	18	VAL
1	E	37	ASP
1	E	51	ASP
1	E	110	SER
1	E	114	ASP
1	E	186	LYS
1	E	187	GLN
1	E	279	SER
2	I	10	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	28	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	99	GLN
1	E	129	ASN
1	E	134(A)	GLN
1	E	135	GLN
1	E	141	ASN
1	E	166	ASN
1	E	300	ASN
2	I	2	HIS
2	I	9	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is 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	STA	I	6	2	10,10,11	0.93	0	9,12,14	0.99	1 (11%)

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	STA	I	6	2	-	0/11/11/12	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	I	6	STA	O-C-CM	-2.39	117.74	125.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	I	6	STA	1	0

## 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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.