



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 3, 2018 – 10:55 AM EST

PDB ID : 6BGP  
Title : Crystal Structure of Human Calpain-3 Protease Core Mutant-C129A  
Authors : Ye, Q.; Campbell, R.L.; Davies, P.L.  
Deposited on : 2017-10-29  
Resolution : 2.75 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	:	rb-20030736
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)	:	rb-20030736

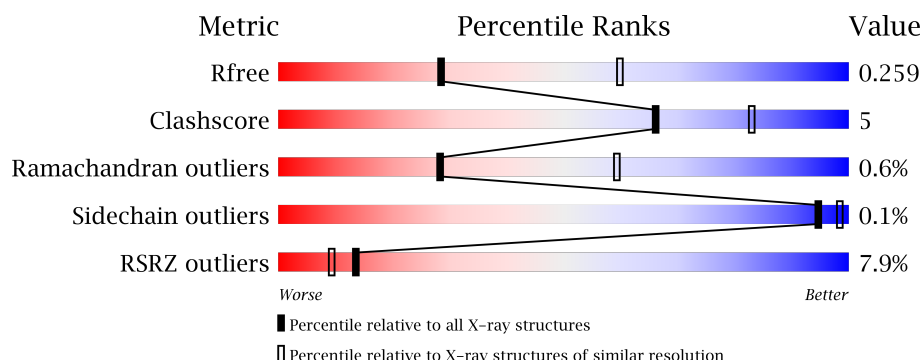
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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	3666 (2.80-2.72)
Clashscore	112137	4174 (2.80-2.72)
Ramachandran outliers	110173	4103 (2.80-2.72)
Sidechain outliers	110143	4106 (2.80-2.72)
RSRZ outliers	101464	3697 (2.80-2.72)

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	382	
1	B	382	
1	C	382	
1	D	382	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10316 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 Calpain-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	0	0	0
			2501	1610	416	461	14			
1	B	314	Total	C	N	O	S	0	0	0
			2588	1665	430	478	15			
1	C	313	Total	C	N	O	S	0	0	0
			2577	1658	430	474	15			
1	D	314	Total	C	N	O	S	0	0	0
			2588	1665	430	478	15			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	129	ALA	CYS	engineered mutation	UNP P20807
A	420	LEU	-	expression tag	UNP P20807
A	421	GLU	-	expression tag	UNP P20807
A	422	HIS	-	expression tag	UNP P20807
A	423	HIS	-	expression tag	UNP P20807
A	424	HIS	-	expression tag	UNP P20807
A	425	HIS	-	expression tag	UNP P20807
A	426	HIS	-	expression tag	UNP P20807
A	427	HIS	-	expression tag	UNP P20807
B	129	ALA	CYS	engineered mutation	UNP P20807
B	420	LEU	-	expression tag	UNP P20807
B	421	GLU	-	expression tag	UNP P20807
B	422	HIS	-	expression tag	UNP P20807
B	423	HIS	-	expression tag	UNP P20807
B	424	HIS	-	expression tag	UNP P20807
B	425	HIS	-	expression tag	UNP P20807
B	426	HIS	-	expression tag	UNP P20807
B	427	HIS	-	expression tag	UNP P20807
C	129	ALA	CYS	engineered mutation	UNP P20807
C	420	LEU	-	expression tag	UNP P20807
C	421	GLU	-	expression tag	UNP P20807

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	422	HIS	-	expression tag	UNP P20807
C	423	HIS	-	expression tag	UNP P20807
C	424	HIS	-	expression tag	UNP P20807
C	425	HIS	-	expression tag	UNP P20807
C	426	HIS	-	expression tag	UNP P20807
C	427	HIS	-	expression tag	UNP P20807
D	129	ALA	CYS	engineered mutation	UNP P20807
D	420	LEU	-	expression tag	UNP P20807
D	421	GLU	-	expression tag	UNP P20807
D	422	HIS	-	expression tag	UNP P20807
D	423	HIS	-	expression tag	UNP P20807
D	424	HIS	-	expression tag	UNP P20807
D	425	HIS	-	expression tag	UNP P20807
D	426	HIS	-	expression tag	UNP P20807
D	427	HIS	-	expression tag	UNP P20807

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

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

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

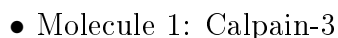
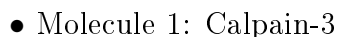
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Cl 1 1	0	0
3	A	1	Total Cl 1 1	0	0
3	D	1	Total Cl 1 1	0	0
3	C	1	Total Cl 1 1	0	0

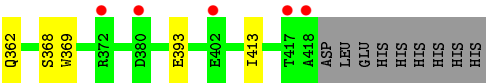
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	14	Total 14	O 14	0	0
4	B	18	Total 18	O 18	0	0
4	C	9	Total 9	O 9	0	0
4	D	9	Total 9	O 9	0	0

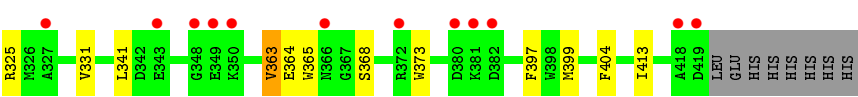
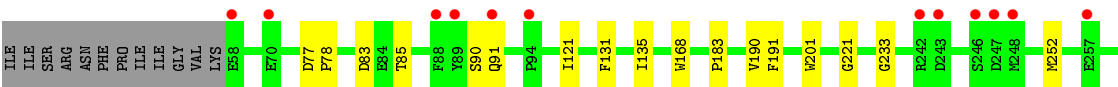


- Molecule 1: Calpain-3





● Molecule 1: Calpain-3



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.10Å 106.18Å 225.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	17.94 – 2.75 17.94 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.5 (17.94-2.75) 100.0 (17.94-2.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.98 (at 2.74Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.209 , 0.258 0.213 , 0.259	Depositor DCC
$R_{free}$ test set	1916 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	62.2	Xtriage
Anisotropy	0.035	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 51.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10316	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.30% 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, CL

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	0.48	0/2573	0.63	0/3480
1	B	0.52	0/2663	0.63	0/3604
1	C	0.44	0/2651	0.64	0/3586
1	D	0.64	0/2663	0.67	0/3604
All	All	0.53	0/10550	0.64	0/14274

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	2501	0	2373	34	0
1	B	2588	0	2460	25	0
1	C	2577	0	2460	16	0
1	D	2588	0	2460	16	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	14	0	0	1	0
4	B	18	0	0	1	0
4	C	9	0	0	0	0
4	D	9	0	0	0	0
All	All	10316	0	9753	91	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 5.

The worst 5 of 91 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:325:ARG:HD2	1:A:332:ARG:H	1.18	1.05
1:D:83:ASP:OD1	1:D:85:THR:HG22	1.69	0.92
1:B:102:PRO:HB2	1:B:103:PRO:HD3	1.57	0.85
1:B:365:TRP:HZ3	1:B:368:SER:O	1.62	0.83
1:A:365:TRP:HZ3	1:A:368:SER:O	1.63	0.81

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	297/382 (78%)	287 (97%)	8 (3%)	2 (1%)	25 56
1	B	310/382 (81%)	299 (96%)	9 (3%)	2 (1%)	28 59
1	C	309/382 (81%)	300 (97%)	8 (3%)	1 (0%)	44 75
1	D	310/382 (81%)	300 (97%)	8 (3%)	2 (1%)	28 59

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1226/1528 (80%)	1186 (97%)	33 (3%)	7 (1%)	28	59

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	221	GLY
1	B	221	GLY
1	C	221	GLY
1	A	363	VAL
1	D	363	VAL

### 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	266/338 (79%)	266 (100%)	0	100	100
1	B	276/338 (82%)	276 (100%)	0	100	100
1	C	275/338 (81%)	275 (100%)	0	100	100
1	D	276/338 (82%)	275 (100%)	1 (0%)	93	97
All	All	1093/1352 (81%)	1092 (100%)	1 (0%)	94	98

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	365	TRP

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

### 5.3.3 RNA ⓘ

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](#)

Of 12 ligands modelled in this entry, 12 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 [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	303/382 (79%)	0.14	16 (5%) 27 21	35, 59, 96, 144	0
1	B	314/382 (82%)	0.35	23 (7%) 16 11	37, 62, 107, 150	0
1	C	313/382 (81%)	0.30	27 (8%) 11 8	37, 65, 114, 142	0
1	D	314/382 (82%)	0.39	32 (10%) 7 5	33, 65, 122, 149	0
All	All	1244/1528 (81%)	0.30	98 (7%) 13 9	33, 62, 111, 150	0

The worst 5 of 98 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	270	THR	6.8
1	D	246	SER	6.7
1	C	271	ASN	6.6
1	B	419	ASP	6.4
1	D	58	GLU	6.1

### 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CA	D	502	1/1	0.84	0.14	-0.90	54,54,54,54	0
2	CA	C	501	1/1	0.97	0.12	-1.01	41,41,41,41	0
2	CA	D	501	1/1	0.97	0.11	-1.07	37,37,37,37	0
3	CL	C	503	1/1	0.97	0.11	-1.21	46,46,46,46	0
3	CL	B	503	1/1	0.98	0.12	-1.58	58,58,58,58	0
2	CA	A	501	1/1	0.98	0.08	-1.72	41,41,41,41	0
2	CA	B	502	1/1	0.97	0.06	-1.99	57,57,57,57	0
2	CA	B	501	1/1	0.99	0.08	-2.02	41,41,41,41	0
3	CL	A	503	1/1	0.97	0.08	-2.25	47,47,47,47	0
2	CA	C	502	1/1	0.98	0.06	-2.82	72,72,72,72	0
3	CL	D	503	1/1	0.99	0.07	-3.69	52,52,52,52	0
2	CA	A	502	1/1	0.98	0.09	-4.39	51,51,51,51	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.