wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 10:32 PM GMT

PDB ID	:	1U5I
Title	:	Crystal Structure analysis of rat m-calpain mutant Lys10 Thr
Authors	:	Hosfield, C.M.; Pal, G.P.; Elce, J.S.; Jia, Z.
Deposited on	:	2004-07-27
Resolution	:	2.86 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

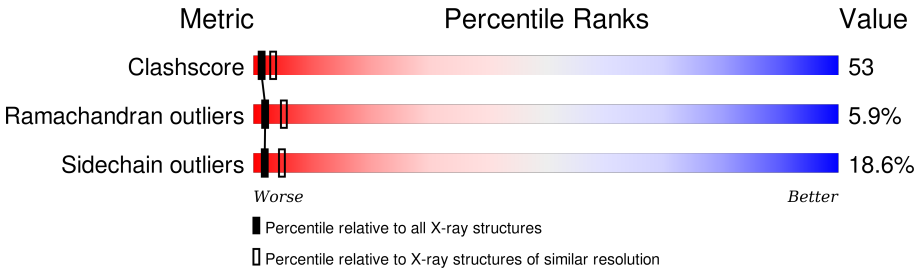
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.86 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	2499 (2.90-2.82)
Ramachandran outliers	100387	2439 (2.90-2.82)
Sidechain outliers	100360	2442 (2.90-2.82)

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 >=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 <=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	700	12% 37% 29% 10% 11%
2	B	184	20% 34% 33% 10% .

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6706 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 2, large [catalytic] subunit precursor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	625	Total	C	N	O	S	0	0	0
			4980	3176	836	945	23			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	10	THR	LYS	ENGINEERED	UNP Q07009
A	105	SER	CYS	ENGINEERED	UNP Q07009

- Molecule 2 is a protein called Calpain small subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	176	Total	C	N	O	S	0	0	0
			1427	897	246	274	10			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	INITIATING METHIONINE	UNP Q64537

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	247	Total	O	0	0
			247	247		
3	B	52	Total	O	0	0
			52	52		

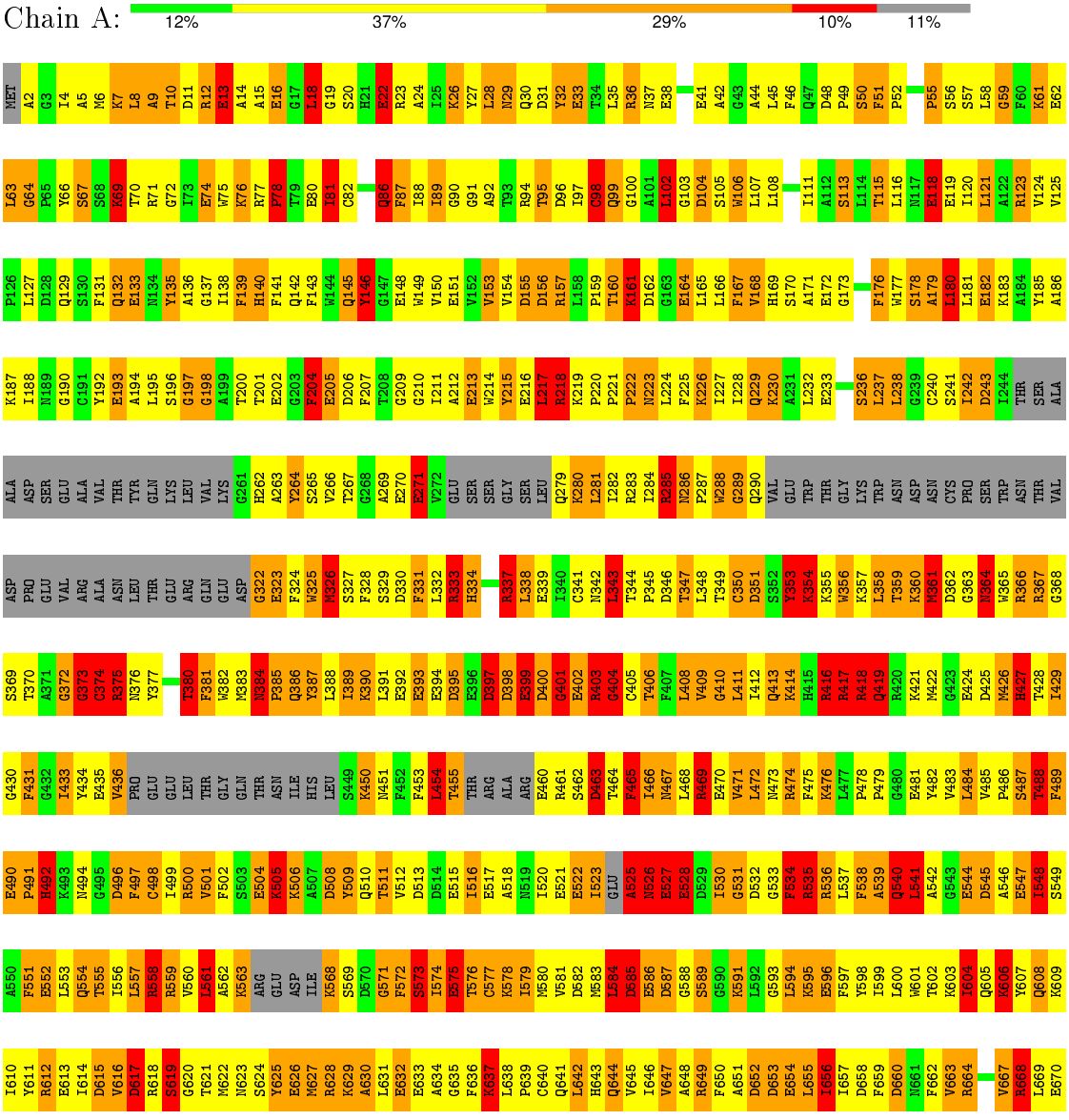
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

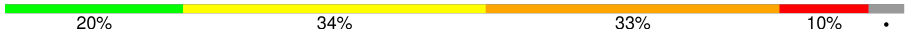
- Molecule 1: Calpain 2, large [catalytic] subunit precursor

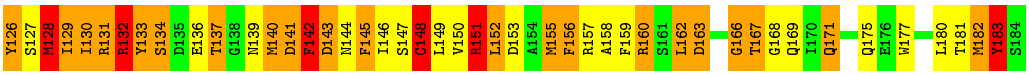
Chain A:



- Molecule 2: Calpain small subunit 1

Chain B:





4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	51.66Å 156.52Å 64.23Å 90.00° 95.39° 90.00°	Depositor
Resolution (Å)	25.00 – 2.86	Depositor
% Data completeness (in resolution range)	90.1 (25.00-2.86)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.191 , 0.283	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6706	wwPDB-VP
Average B, all atoms (Å^2)	59.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	3.07	491/5083 (9.7%)	2.24	221/6854 (3.2%)
2	B	2.76	104/1454 (7.2%)	2.06	43/1955 (2.2%)
All	All	3.01	595/6537 (9.1%)	2.20	264/8809 (3.0%)

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	19
2	B	0	2
All	All	0	21

The worst 5 of 595 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	400	ASP	C-N	31.66	1.90	1.33
1	A	402	GLU	C-N	23.12	1.87	1.34
2	B	133	TYR	CD2-CE2	15.60	1.62	1.39
1	A	417	ARG	CG-CD	14.55	1.88	1.51
1	A	628	ARG	CZ-NH2	14.16	1.51	1.33

The worst 5 of 264 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	400	ASP	C-N-CA	-23.66	72.61	122.30
1	A	23	ARG	NE-CZ-NH1	22.54	131.57	120.30
1	A	402	GLU	CA-C-N	-20.02	73.17	117.20
1	A	403	ARG	CA-C-N	-19.27	77.66	116.20
1	A	23	ARG	NE-CZ-NH2	-18.54	111.03	120.30

There are no chirality outliers.

5 of 21 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	146	TYR	Sidechain
1	A	18	LEU	Mainchain
1	A	197	GLY	Peptide
1	A	9	ALA	Mainchain
1	A	98	CYS	Mainchain

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	4980	0	4809	519	10
2	B	1427	0	1375	156	9
3	A	247	0	0	67	0
3	B	52	0	0	23	0
All	All	6706	0	6184	668	10

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 53.

The worst 5 of 668 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:218:ARG:CG	1:A:218:ARG:CB	1.74	1.66
1:A:409:VAL:CG1	1:A:409:VAL:CB	1.76	1.64
1:A:628:ARG:CG	1:A:628:ARG:CD	1.76	1.63
1:A:699:VAL:CG1	1:A:699:VAL:CB	1.77	1.62
1:A:455:THR:CG2	1:A:455:THR:CB	1.80	1.60

The worst 5 of 10 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:A:527:GLU:OE1	2:B:19:LYS:CE[1_554]	0.67	1.53
1:A:527:GLU:CD	2:B:19:LYS:CE[1_554]	0.84	1.36

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:527:GLU:CD	2:B:19:LYS:NZ[1_554]	1.15	1.05
1:A:527:GLU:OE1	2:B:19:LYS:CD[1_554]	1.29	0.91
1:A:527:GLU:OE2	2:B:19:LYS:NZ[1_554]	1.33	0.87

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	609/700 (87%)	506 (83%)	70 (12%)	33 (5%)	2 6
2	B	174/184 (95%)	130 (75%)	31 (18%)	13 (8%)	1 3
All	All	783/884 (89%)	636 (81%)	101 (13%)	46 (6%)	2 5

5 of 46 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	133	GLU
1	A	351	ASP
1	A	354	LYS
1	A	373	GLY
1	A	399	GLU

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	522/603 (87%)	424 (81%)	98 (19%)	2 5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	B	155/162 (96%)	127 (82%)	28 (18%)	2 5
All	All	677/765 (88%)	551 (81%)	126 (19%)	2 5

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	419	GLN
1	A	535	ARG
2	B	123	GLN
1	A	450	LYS
1	A	484	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	379	ASN
1	A	413	GLN
2	B	47	HIS
1	A	384	ASN
1	A	386	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

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

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

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

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

6.3 Carbohydrates

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

6.4 Ligands

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

6.5 Other polymers

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