wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 04:46 AM EDT

PDB ID	:	1KFU
Title	:	Crystal Structure of Human m-Calpain Form II
Authors	:	Strobl, S.; Fernandez-Catalan, C.; Braun, M.; Huber, R.; Masumoto, H.; Nakagawa, K.; Irie, A.; Sorimachi, H.; Bourenkow, G.; Bartunik, H.; Suzuki, K.; Bode, W.
Deposited on	:	2001-11-23
Resolution	:	2.50 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

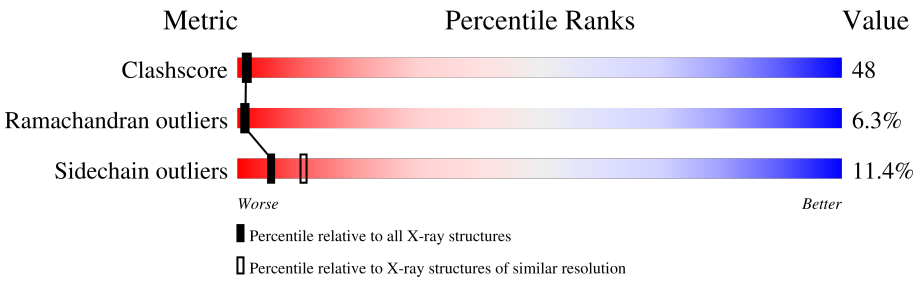
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.50 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	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 >=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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	L	699	34% 52% 12% •
2	S	184	40% 53% 7% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7530 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 M-CALPAIN LARGE SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	L	699	5628	3568	955	1079	26	0	0	0

- Molecule 2 is a protein called M-CALPAIN SMALL SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	S	184	1491	935	255	290	11	10	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	L	287	287	287	0	0
3	S	124	124	124	0	0

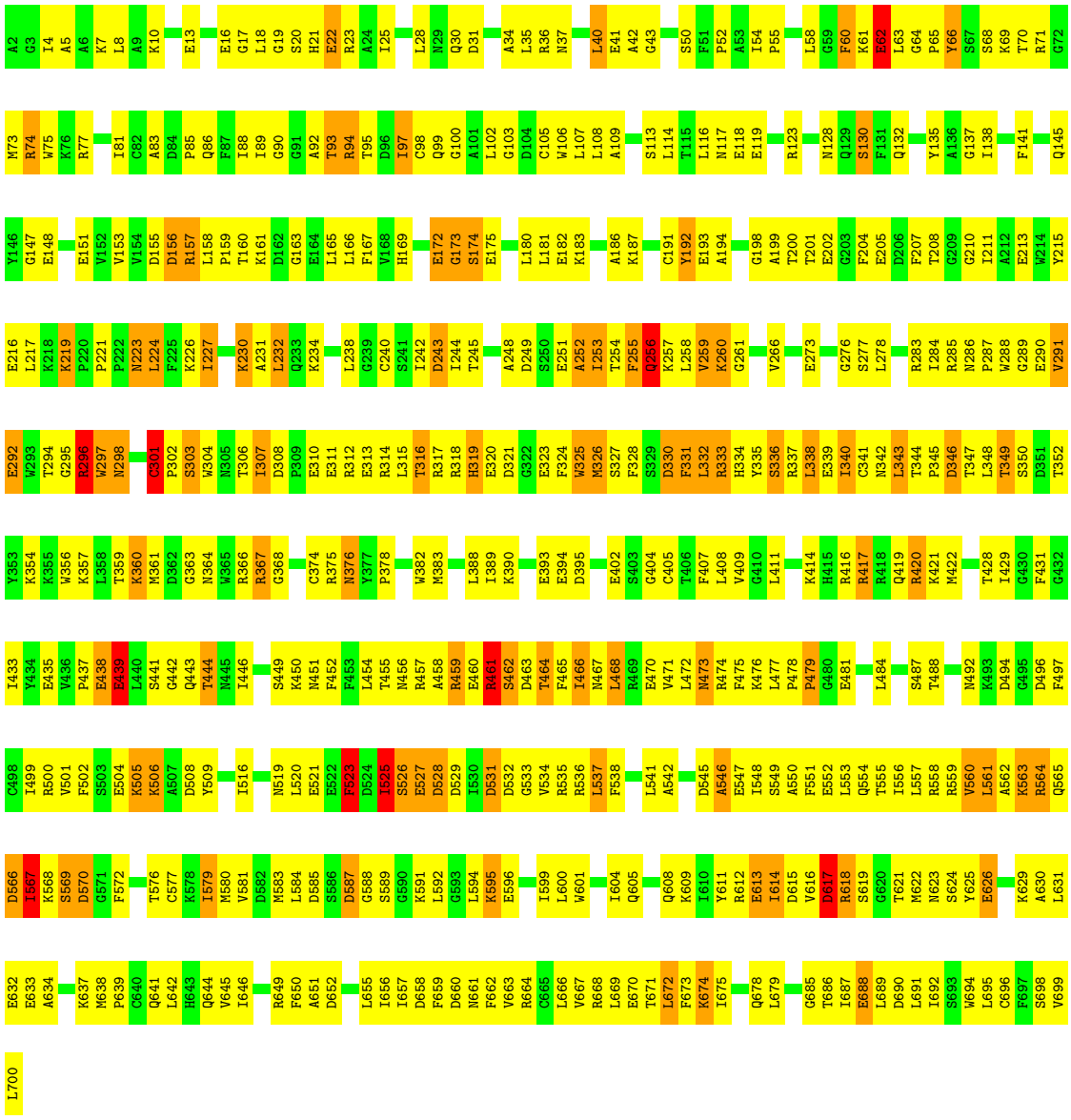
3 Residue-property plots (i)

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. 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. 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.

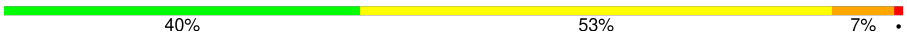
- Molecule 1: M-CALPAIN LARGE SUBUNIT

Chain L: 34% 52% 12%



- Molecule 2: M-CALPAIN SMALL SUBUNIT

Chain S:



T785 H786 N789 I790 E791 A792 N793 E794 S795 E796 E797 V798 R799 Q800 F801 R802 R803 L804 F805 A806 Q807 G810 D811 D812 M813 E814 V815 S816 E819 L820 M821 K826 V827 V828 T829 R830 H831 P832 D833 L834 K835 T836 D837 G840 C844 R845 S846 V850 M851 D852 S853 D854 T855

L859 G860 F861 E862 E863 F864 K865 Y866 L867 W868 N869 N870 I871 K872 R873 W874 Q875 A876 I877 Y878 K879 D882 T883 D884 R885 S886 G887 T888 I889 C890 E893 L894 P895 F898 A901 G902 F903 H904 L905 N906 E907 H908 L909 Y910 N911 M912 I913 I914 R915 R916 Y917 S918 D919 E920 S921

G922 D925 F926 L936 D937 A938 M939 A942 F943 L946 D947 K948 T951 G952 Q953 I954 Q955 V956 N957 I958 Q959 E960 W961 L962 Q963 L964 T965 M966 Y967 S968

4 Data and refinement statistics (i)

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, α, β, γ	51.88Å 169.84Å 64.44Å 90.00° 95.12° 90.00°	Depositor
Resolution (Å)	30.00 – 2.50	Depositor
% Data completeness (in resolution range)	(Not available) (30.00-2.50)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.221 , 0.276	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7530	wwPDB-VP
Average B, all atoms (Å^2)	17.0	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	L	0.64	5/5752 (0.1%)	0.83	8/7768 (0.1%)
2	S	0.58	2/1520 (0.1%)	0.73	1/2046 (0.0%)
All	All	0.63	7/7272 (0.1%)	0.81	9/9814 (0.1%)

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	696	CYS	CB-SG	8.07	1.96	1.82
1	L	301	CYS	CB-SG	-7.64	1.69	1.82
2	S	968	SER	C-O	6.41	1.35	1.23
2	S	968	SER	C-OXT	6.00	1.34	1.23
1	L	62	GLU	C-N	-5.92	1.20	1.34

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	S	968	SER	CA-C-O	7.14	135.09	120.10
1	L	459	ARG	NE-CZ-NH2	6.75	123.67	120.30
1	L	62	GLU	C-N-CA	6.71	138.47	121.70
1	L	461	ARG	NE-CZ-NH2	6.37	123.49	120.30
1	L	326	MET	CG-SD-CE	6.25	110.20	100.20

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	L	5628	0	5466	563	0
2	S	1491	0	1423	135	0
3	L	287	0	0	21	0
3	S	124	0	0	10	0
All	All	7530	0	6889	668	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 48.

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:L:248:ALA:O	1:L:259:VAL:HG12	1.36	1.21
1:L:525:ILE:HG22	1:L:595:LYS:HD2	1.27	1.17
1:L:61:LYS:H	1:L:64:GLY:HA3	1.05	1.11
2:S:869:ASN:HA	2:S:872:LYS:HE2	1.24	1.10
1:L:74:ARG:H	1:L:74:ARG:HD2	1.10	1.09

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	L	697/699 (100%)	546 (78%)	103 (15%)	48 (7%)	1 1
2	S	182/184 (99%)	157 (86%)	18 (10%)	7 (4%)	3 4
All	All	879/883 (100%)	703 (80%)	121 (14%)	55 (6%)	1 1

5 of 55 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	253	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	L	256	GLN
1	L	296	ARG
1	L	298	ASN
1	L	301	CYS

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	L	602/602 (100%)	528 (88%)	74 (12%)	4 9
2	S	162/162 (100%)	149 (92%)	13 (8%)	12 23
All	All	764/764 (100%)	677 (89%)	87 (11%)	5 11

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	555	THR
1	L	672	LEU
1	L	566	ASP
1	L	595	LYS
2	S	789	ASN

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

Mol	Chain	Res	Type
1	L	605	GLN
2	S	789	ASN
1	L	661	ASN
2	S	793	ASN
1	L	223	ASN

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)

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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	L	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	L	62:GLU	C	63:LEU	N	1.20

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.