



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 26, 2017 – 04:36 PM EDT

PDB ID : 3LOF
Title : C-terminal domain of human heat shock 70kDa protein 1B.
Authors : Osipiuk, J.; Gu, M.; Mihelic, M.; Orton, K.; Morimoto, R.I.; Joachimiak, A.;
Midwest Center for Structural Genomics (MCSG)
Deposited on : unknown
Resolution : 2.40 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Xtrriage (Phenix) : 1.9-1692
EDS : rb-20030345
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-20030345

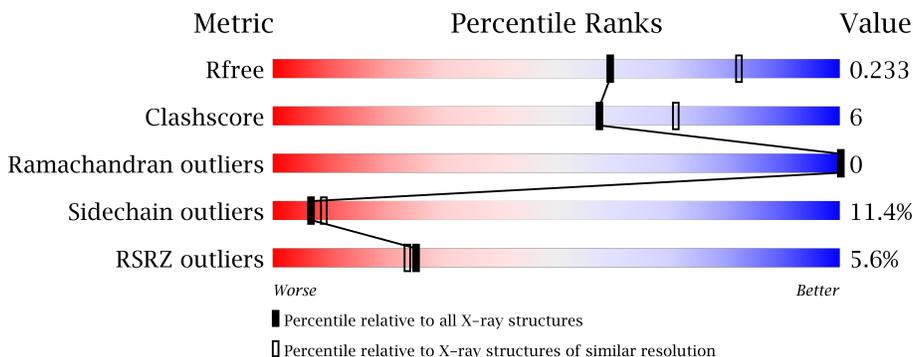
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (2.40-2.40)

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 $\leq 5\%$. 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	113	
1	B	113	
1	C	113	
1	D	113	
1	E	113	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	113	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into four segments: a small red segment at the beginning labeled '3%', followed by a large green segment labeled '59%', a small yellow segment labeled '9%', and a grey segment at the end labeled '29%'. A small black dot is located on the grey segment.</p>

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3934 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 Heat shock 70 kDa protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	84	Total 659	C 410	N 112	O 134	S 2	Se 1	0	1	0
1	B	80	Total 622	C 386	N 106	O 127	S 2	Se 1	0	0	0
1	C	80	Total 628	C 391	N 107	O 127	S 2	Se 1	0	1	0
1	D	79	Total 623	C 388	N 106	O 126	S 2	Se 1	0	1	0
1	E	82	Total 641	C 398	N 109	O 131	S 2	Se 1	0	1	0
1	F	80	Total 627	C 390	N 107	O 127	S 2	Se 1	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	529	SER	-	expression tag	UNP P08107
A	530	ASN	-	expression tag	UNP P08107
A	531	ALA	-	expression tag	UNP P08107
A	532	ALA	-	expression tag	UNP P08107
A	533	ALA	-	expression tag	UNP P08107
B	529	SER	-	expression tag	UNP P08107
B	530	ASN	-	expression tag	UNP P08107
B	531	ALA	-	expression tag	UNP P08107
B	532	ALA	-	expression tag	UNP P08107
B	533	ALA	-	expression tag	UNP P08107
C	529	SER	-	expression tag	UNP P08107
C	530	ASN	-	expression tag	UNP P08107
C	531	ALA	-	expression tag	UNP P08107
C	532	ALA	-	expression tag	UNP P08107
C	533	ALA	-	expression tag	UNP P08107
D	529	SER	-	expression tag	UNP P08107
D	530	ASN	-	expression tag	UNP P08107

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	531	ALA	-	expression tag	UNP P08107
D	532	ALA	-	expression tag	UNP P08107
D	533	ALA	-	expression tag	UNP P08107
E	529	SER	-	expression tag	UNP P08107
E	530	ASN	-	expression tag	UNP P08107
E	531	ALA	-	expression tag	UNP P08107
E	532	ALA	-	expression tag	UNP P08107
E	533	ALA	-	expression tag	UNP P08107
F	529	SER	-	expression tag	UNP P08107
F	530	ASN	-	expression tag	UNP P08107
F	531	ALA	-	expression tag	UNP P08107
F	532	ALA	-	expression tag	UNP P08107
F	533	ALA	-	expression tag	UNP P08107

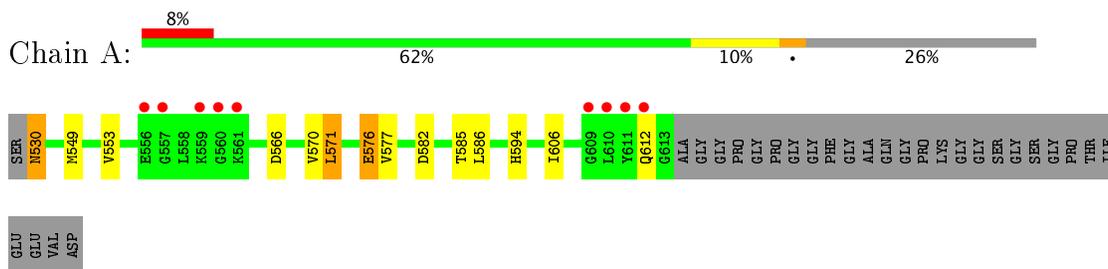
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	28	Total O 28 28	0	0
2	B	14	Total O 14 14	0	0
2	C	26	Total O 26 26	0	0
2	D	31	Total O 31 31	0	0
2	E	24	Total O 24 24	0	0
2	F	11	Total O 11 11	0	0

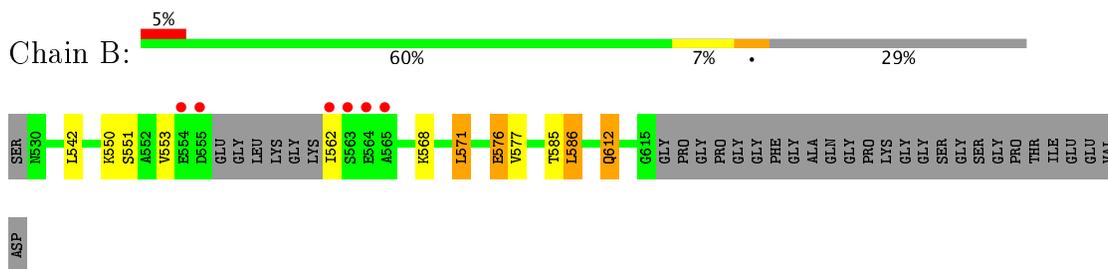
3 Residue-property plots [i](#)

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

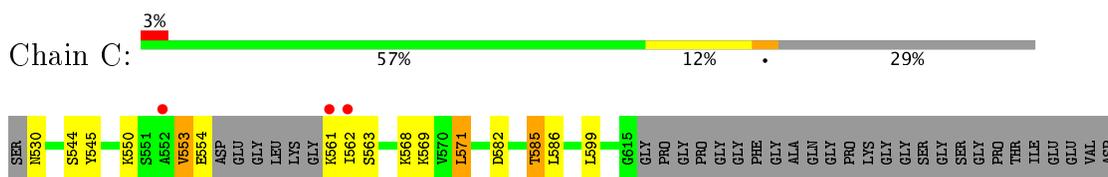
- Molecule 1: Heat shock 70 kDa protein 1



- Molecule 1: Heat shock 70 kDa protein 1



- Molecule 1: Heat shock 70 kDa protein 1



- Molecule 1: Heat shock 70 kDa protein 1

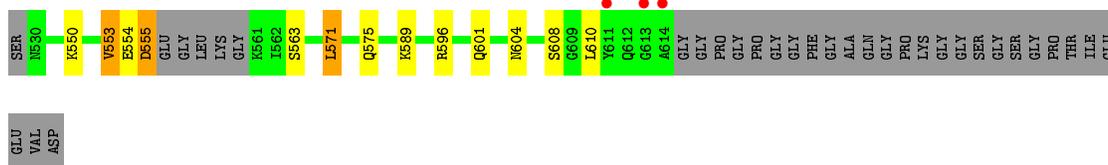


- Molecule 1: Heat shock 70 kDa protein 1





• Molecule 1: Heat shock 70 kDa protein 1



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	70.70Å 71.88Å 143.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.00 – 2.40 32.04 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.0 (32.00-2.40) 99.0 (32.04-2.40)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.51 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.192 , 0.235 0.188 , 0.233	Depositor DCC
R_{free} test set	1475 reflections (5.36%)	DCC
Wilson B-factor (Å ²)	39.2	Xtrriage
Anisotropy	0.070	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 50.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.023 for k,h,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3934	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.13% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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](#)

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	0.97	0/668	0.80	0/892
1	B	0.97	2/627 (0.3%)	0.81	0/838
1	C	1.05	0/636	0.82	0/849
1	D	0.95	0/631	0.90	3/842 (0.4%)
1	E	0.93	0/649	0.81	0/866
1	F	0.93	0/632	0.84	1/844 (0.1%)
All	All	0.97	2/3843 (0.1%)	0.83	4/5131 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	612	GLN	CB-CG	-5.05	1.39	1.52
1	B	576	GLU	CG-CD	5.01	1.59	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	582	ASP	CB-CG-OD2	-5.76	113.12	118.30
1	D	586	LEU	CA-CB-CG	-5.72	102.15	115.30
1	F	596	ARG	NE-CZ-NH2	5.32	122.96	120.30
1	D	582	ASP	CB-CG-OD1	5.07	122.86	118.30

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	659	0	657	10	0
1	B	622	0	610	3	0
1	C	628	0	623	9	0
1	D	623	0	618	6	0
1	E	641	0	632	10	0
1	F	627	0	620	5	0
2	A	28	0	0	2	0
2	B	14	0	0	0	0
2	C	26	0	0	3	0
2	D	31	0	0	2	0
2	E	24	0	0	1	0
2	F	11	0	0	1	0
All	All	3934	0	3760	42	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 6.

The worst 5 of 42 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:C:585:THR:CG2	2:C:33:HOH:O	2.22	0.87
1:C:553:VAL:HG11	1:C:571:LEU:HD13	1.70	0.71
1:A:553:VAL:HG21	1:A:571:LEU:HD13	1.72	0.68
1:B:553:VAL:HG11	1:B:571:LEU:HD13	1.75	0.68
1:D:585:THR:CG2	2:D:89:HOH:O	2.43	0.67

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	83/113 (74%)	82 (99%)	1 (1%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	76/113 (67%)	75 (99%)	1 (1%)	0	100	100
1	C	77/113 (68%)	74 (96%)	3 (4%)	0	100	100
1	D	76/113 (67%)	76 (100%)	0	0	100	100
1	E	79/113 (70%)	76 (96%)	3 (4%)	0	100	100
1	F	76/113 (67%)	71 (93%)	5 (7%)	0	100	100
All	All	467/678 (69%)	454 (97%)	13 (3%)	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	71/85 (84%)	67 (94%)	4 (6%)	25	39
1	B	66/85 (78%)	58 (88%)	8 (12%)	6	7
1	C	67/85 (79%)	56 (84%)	11 (16%)	2	3
1	D	67/85 (79%)	60 (90%)	7 (10%)	8	12
1	E	68/85 (80%)	60 (88%)	8 (12%)	6	8
1	F	67/85 (79%)	59 (88%)	8 (12%)	6	8
All	All	406/510 (80%)	360 (89%)	46 (11%)	7	9

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	569	LYS
1	D	571	LEU
1	F	571	LEU
1	C	571	LEU
1	D	550	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	540	ASN
1	C	604	ASN
1	D	604	ASN
1	B	612	GLN
1	D	601	GLN

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 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 [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, 95th 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 ²)	Q<0.9
1	A	83/113 (73%)	0.23	9 (10%) 6 6	15, 28, 46, 50	0
1	B	79/113 (69%)	0.16	6 (7%) 15 13	15, 31, 54, 62	0
1	C	79/113 (69%)	-0.03	3 (3%) 41 40	10, 25, 45, 53	0
1	D	78/113 (69%)	-0.15	0 100 100	14, 22, 50, 58	0
1	E	81/113 (71%)	0.24	6 (7%) 15 14	13, 27, 46, 57	0
1	F	79/113 (69%)	0.05	3 (3%) 41 40	12, 29, 47, 50	0
All	All	479/678 (70%)	0.09	27 (5%) 25 24	10, 27, 49, 62	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	610	LEU	5.2
1	C	562	ILE	5.1
1	A	560	GLY	4.8
1	C	561	LYS	4.2
1	A	559	LYS	4.0

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 carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.