



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:21 AM GMT

PDB ID : 3AGY
Title : Crystal structure of human Hsp40 Hdj1 peptide-binding domain complexed with a C-terminal peptide of Hsp70
Authors : Suzuki, H.; Noguchi, S.; Satow, Y.
Deposited on : 2010-04-12
Resolution : 1.85 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
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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

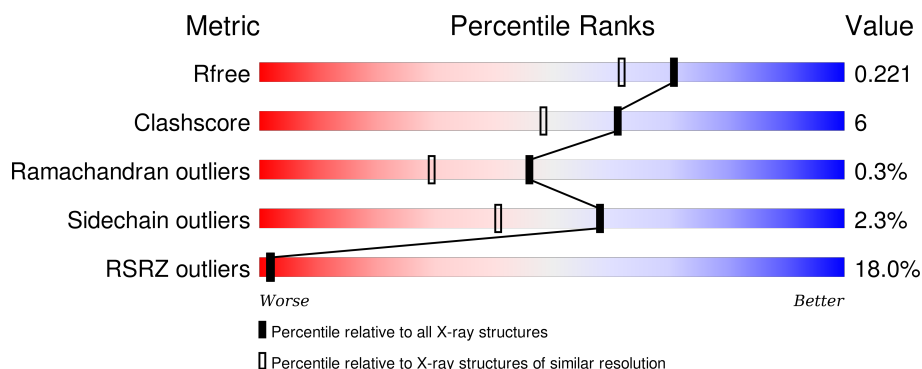
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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}	91344	1745 (1.86-1.86)
Clashscore	102246	1898 (1.86-1.86)
Ramachandran outliers	100387	1875 (1.86-1.86)
Sidechain outliers	100360	1875 (1.86-1.86)
RSRZ outliers	91569	1747 (1.86-1.86)

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	181	
1	B	181	
2	C	8	
2	D	8	
2	F	8	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3220 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 DnaJ homolog subfamily B member 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	176	Total	C	N	O	S	0	2	0
			1406	894	250	257	5			
1	B	176	Total	C	N	O	S	0	7	0
			1425	905	258	257	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	160	MET	-	EXPRESSION TAG	UNP P25685
B	160	MET	-	EXPRESSION TAG	UNP P25685

- Molecule 2 is a protein called peptide of Heat shock cognate 71 kDa protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	7	Total	C	N	O	0	0	0
			51	32	7	12			
2	D	8	Total	C	N	O	0	0	0
			60	36	8	16			
2	F	6	Total	C	N	O	0	0	0
			47	30	6	11			

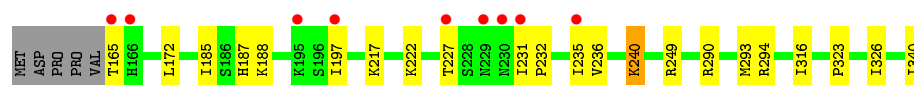
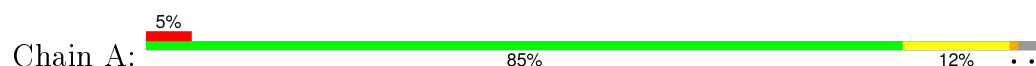
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	123	Total	O	0	0
			123	123		
3	B	104	Total	O	0	0
			104	104		
3	F	4	Total	O	0	0
			4	4		

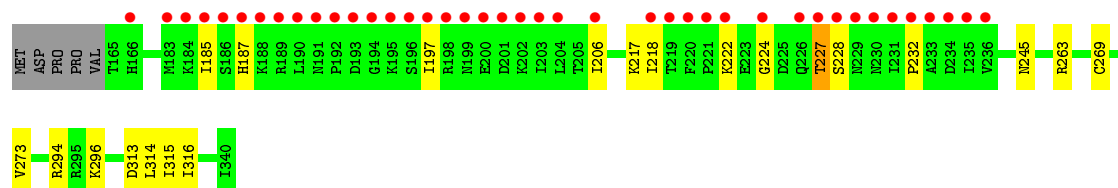
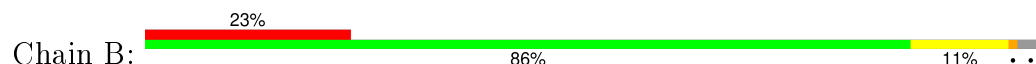
3 Residue-property plots [i](#)

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.

- Molecule 1: DnaJ homolog subfamily B member 1



- Molecule 1: DnaJ homolog subfamily B member 1



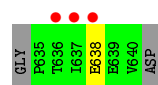
- Molecule 2: peptide of Heat shock cognate 71 kDa protein



- Molecule 2: peptide of Heat shock cognate 71 kDa protein



- Molecule 2: peptide of Heat shock cognate 71 kDa protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	104.75Å 41.13Å 62.87Å 90.00° 97.42° 90.00°	Depositor
Resolution (Å)	42.71 – 1.85 42.71 – 1.85	Depositor EDS
% Data completeness (in resolution range)	98.9 (42.71-1.85) 98.8 (42.71-1.85)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.65 (at 1.86Å)	Xtriage
Refinement program	REFMAC 5.5.0089	Depositor
R, R_{free}	0.225 , 0.247 0.225 , 0.221	Depositor DCC
R_{free} test set	2282 reflections (5.30%)	DCC
Wilson B-factor (Å ²)	24.9	Xtriage
Anisotropy	0.300	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 48.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 45424 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3220	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.52% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.67	0/1443	0.68	0/1947
1	B	0.73	1/1491 (0.1%)	0.75	1/2006 (0.0%)
2	C	0.58	0/51	0.48	0/69
2	D	0.58	0/60	0.55	0/80
2	F	0.41	0/47	0.52	0/63
All	All	0.69	1/3092 (0.0%)	0.71	1/4165 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	269	CYS	CB-SG	-9.27	1.66	1.82

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	263	ARG	NE-CZ-NH1	-8.09	116.25	120.30

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	1406	0	1470	22	0
1	B	1425	0	1507	11	0
2	C	51	0	48	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	60	0	52	3	0
2	F	47	0	46	1	0
3	A	123	0	0	1	0
3	B	104	0	0	0	0
3	F	4	0	0	0	0
All	All	3220	0	3123	34	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.

All (34) 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:290:ARG:H	1:A:293[A]:MET:HE3	1.00	1.10
1:A:290:ARG:N	1:A:293[A]:MET:HE3	1.78	0.97
1:A:290:ARG:H	1:A:293[A]:MET:CE	1.88	0.85
1:A:185:ILE:HD13	1:A:235:ILE:HD13	1.65	0.79
1:A:185:ILE:HD13	1:A:235:ILE:CD1	2.22	0.70
1:A:217:LYS:HE3	1:A:236:VAL:HG11	1.74	0.69
1:A:290:ARG:CZ	1:A:293[A]:MET:HE2	2.25	0.66
1:A:185:ILE:CD1	1:A:235:ILE:HD13	2.26	0.65
1:A:188:LYS:HB3	1:A:197:ILE:HG21	1.79	0.65
1:B:206:ILE:HG21	1:B:218:ILE:HD13	1.85	0.59
1:A:185:ILE:CD1	1:A:235:ILE:CD1	2.83	0.56
1:A:290:ARG:NH2	1:A:293[A]:MET:HE1	2.24	0.53
1:A:340:ILE:HG23	3:A:370:HOH:O	2.08	0.52
1:B:185:ILE:HG22	2:D:637:ILE:HG12	1.93	0.50
2:D:640:VAL:HG22	2:D:641:ASP:H	1.75	0.50
1:A:290:ARG:CZ	1:A:293[A]:MET:CE	2.88	0.50
1:A:185:ILE:HD12	2:C:637:ILE:HG12	1.93	0.50
1:A:172:LEU:HD21	1:A:249:ARG:HB2	1.94	0.49
1:B:217:LYS:HB2	2:F:638:GLU:HB3	1.95	0.48
1:B:294:ARG:HA	1:B:316:ILE:O	2.14	0.47
1:A:294:ARG:HA	1:A:316:ILE:O	2.14	0.47
1:A:323:PRO:HG2	1:A:326:ILE:HD11	1.96	0.46
1:A:188:LYS:HB3	1:A:197:ILE:CG2	2.45	0.46
1:A:240:LYS:HD2	1:A:240:LYS:HA	1.77	0.46
1:B:224:GLY:HA2	2:D:635:PRO:HG3	2.01	0.43
1:B:273:VAL:HG21	1:B:316:ILE:HD11	2.01	0.42
1:B:296[B]:LYS:HD2	1:B:313:ASP:OD2	2.19	0.42
1:B:314:LEU:HD23	1:B:315:ILE:N	2.34	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:222:LYS:HD2	1:A:231:ILE:HD11	2.01	0.41
1:B:197:ILE:HG12	1:B:227:THR:HG22	2.02	0.41
1:A:222:LYS:HB3	1:A:232:PRO:HG2	2.02	0.41
1:B:314:LEU:C	1:B:314:LEU:HD23	2.41	0.41
1:B:222:LYS:HB3	1:B:232:PRO:HG2	2.03	0.41
1:A:290:ARG:NH2	1:A:293[A]:MET:CE	2.84	0.40

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	176/181 (97%)	175 (99%)	1 (1%)	0	100	100
1	B	181/181 (100%)	177 (98%)	3 (2%)	1 (1%)	30	13
2	C	5/8 (62%)	5 (100%)	0	0	100	100
2	D	6/8 (75%)	5 (83%)	1 (17%)	0	100	100
2	F	4/8 (50%)	4 (100%)	0	0	100	100
All	All	372/386 (96%)	366 (98%)	5 (1%)	1 (0%)	46	29

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	228	SER

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	163/166 (98%)	159 (98%)	4 (2%)	55	37
1	B	168/166 (101%)	164 (98%)	4 (2%)	57	39
2	C	6/7 (86%)	6 (100%)	0	100	100
2	D	7/7 (100%)	6 (86%)	1 (14%)	4	0
2	F	6/7 (86%)	6 (100%)	0	100	100
All	All	350/353 (99%)	341 (97%)	9 (3%)	58	36

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

Mol	Chain	Res	Type
1	A	165	THR
1	A	187	HIS
1	A	227	THR
1	A	240	LYS
1	B	187	HIS
1	B	227	THR
1	B	245[A]	ASN
1	B	245[B]	ASN
2	D	636	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	336	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 [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

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	176/181 (97%)	0.45	9 (5%) 32 30	13, 29, 52, 64	0
1	B	176/181 (97%)	1.26	41 (23%) 1 1	12, 32, 88, 90	0
2	C	7/8 (87%)	2.74	6 (85%) 0 0	71, 71, 74, 74	0
2	D	8/8 (100%)	4.43	8 (100%) 0 0	102, 102, 103, 103	0
2	F	6/8 (75%)	2.09	3 (50%) 0 0	65, 66, 67, 68	0
All	All	373/386 (96%)	0.99	67 (17%) 2 2	12, 32, 87, 103	0

All (67) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	231	ILE	13.0
1	B	227	THR	8.7
1	B	229	ASN	8.0
1	B	230	ASN	7.5
1	B	197	ILE	7.4
1	B	193	ASP	6.4
1	B	196	SER	6.2
1	B	228	SER	6.0
2	D	640	VAL	5.6
1	A	231	ILE	5.3
2	D	636	THR	5.3
1	B	190	LEU	5.0
1	B	232	PRO	5.0
2	D	635	PRO	5.0
1	A	166	HIS	4.7
2	D	638	GLU	4.6
2	D	641	ASP	4.6
1	A	197	ILE	4.6
1	B	200	GLU	4.5
2	C	634	GLY	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	189	ARG	4.3
1	B	233	ALA	4.3
1	B	203	ILE	4.2
2	D	637	ILE	4.2
1	B	195[A]	LYS	4.1
1	B	166	HIS	4.0
1	B	204	LEU	3.9
1	B	183	MET	3.9
1	A	229	ASN	3.8
1	B	236	VAL	3.7
1	B	199	ASN	3.7
1	B	222	LYS	3.5
1	B	185	ILE	3.4
1	B	194	GLY	3.4
1	B	202	LYS	3.3
1	B	192	PRO	3.2
1	B	226	GLN	3.2
2	F	637	ILE	3.2
1	B	201	ASP	3.2
1	B	184	LYS	3.2
2	D	634	GLY	3.2
2	D	639	GLU	3.1
2	F	636	THR	3.1
1	B	221	PRO	3.1
1	B	187	HIS	3.0
1	B	220	PHE	2.9
1	A	230	ASN	2.9
1	A	227	THR	2.9
1	B	191	ASN	2.9
2	C	636	THR	2.8
1	B	235	ILE	2.8
1	B	186	SER	2.8
2	C	635	PRO	2.8
1	B	206	ILE	2.8
2	C	638	GLU	2.6
1	B	224	GLY	2.6
1	B	218	ILE	2.5
1	B	188	LYS	2.4
2	C	640	VAL	2.4
2	C	637	ILE	2.4
1	B	219	THR	2.4
1	B	234	ASP	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	235	ILE	2.3
1	B	198	ARG	2.2
2	F	638	GLU	2.2
1	A	165	THR	2.1
1	A	195	LYS	2.0

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 [i](#)

There are no such residues in this entry.