wwPDB NMR Structure Validation Summary Report

May 7, 2024 – 12:33 pm BST

PDB ID	:	4AXP
BMRB ID	:	18523
Title	:	NMR structure of Hsp12, a protein induced by and required for dietary restriction-induced lifespan extension in yeast.
Authors	:	Herbert, A.P.; Riesen, M.; Bloxam, L.; Kosmidou, E.; Wareing, B.M.; Johnson, J.R.; Phelan, M.M.; Pennington, S.R.; Lian, L.Y.; Morgan, A.
Deposited on	:	2012-06-13

This is a wwPDB NMR 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/NMRValidationReportHelp
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:

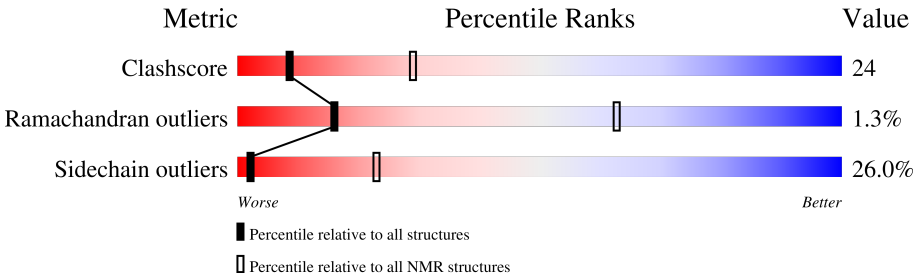
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 91%.

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)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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%

Mol	Chain	Length	Quality of chain
1	A	109	32% 17% 7% 43%

2 Ensemble composition and analysis

This entry contains 20 models. Model 2 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:23-A:34 (12)	0.10	5
2	A:46-A:67 (22)	1.16	12
3	A:74-A:101 (28)	0.64	2

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 4 clusters. No single-model clusters were found.

Cluster number	Models
1	2, 4, 5, 7, 8, 15, 19
2	1, 9, 10, 12, 20
3	6, 11, 14, 17, 18
4	3, 13, 16

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1439 atoms, of which 619 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called 12 KDA HEAT SHOCK PROTEIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	109	1439	496	619	146	176	2	0

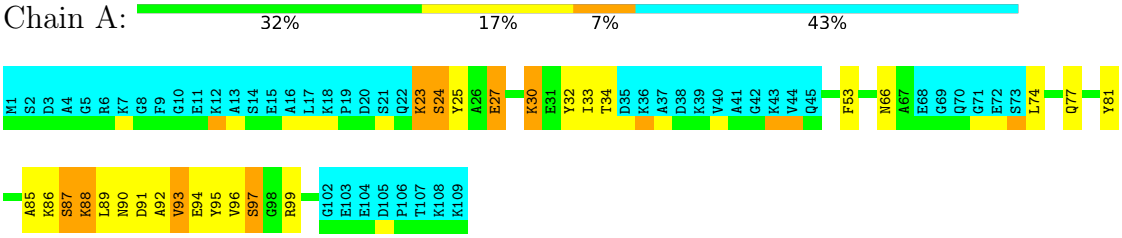
4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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 outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

• Molecule 1: 12 KDA HEAT SHOCK PROTEIN

Chain A:

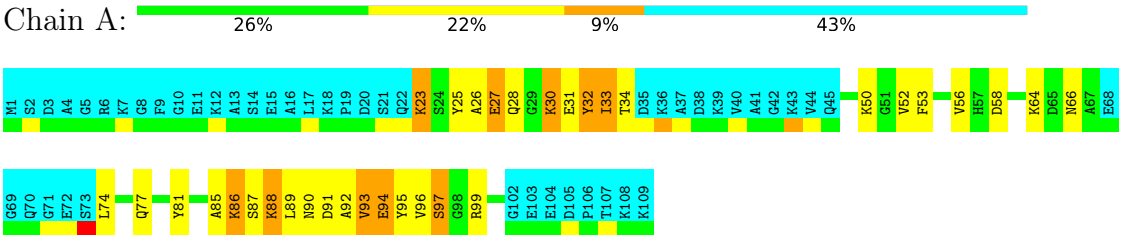


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 2. Colouring as in section 4.1 above.

• Molecule 1: 12 KDA HEAT SHOCK PROTEIN

Chain A:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *NONE*.

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CYANA	refinement	
CYANA	structure solution	2.1
CcpNmr Analysis	structure solution	2.1
TopSpin	structure solution	2.1
DANGLE	structure solution	1

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1219
Number of shifts mapped to atoms	1060
Number of unparsed shifts	0
Number of shifts with mapping errors	159
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	91%

6 Model quality

6.1 Standard geometry

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	477	367	451	22±3
All	All	9540	7340	9020	441

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 24.

5 of 58 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:88:LYS:O	1:A:92:ALA:HB3	0.98	1.59	19	19
1:A:30:LYS:O	1:A:34:THR:HG22	0.93	1.64	10	20
1:A:92:ALA:O	1:A:96:VAL:N	0.70	2.24	10	1
1:A:74:LEU:C	1:A:74:LEU:HD22	0.70	2.07	12	2
1:A:90:ASN:O	1:A:94:GLU:N	0.70	2.24	20	15

6.3 Torsion angles

6.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. 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	62/109 (57%)	51±1 (82±2%)	10±2 (16±3%)	1±1 (1±1%)	16	63
All	All	1240/2180 (57%)	1022 (82%)	202 (16%)	16 (1%)	16	63

All 2 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	24	SER	12
1	A	23	LYS	4

6.3.2 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 PDB entries followed by that with respect to all NMR entries. 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	48/83 (58%)	36±3 (74±6%)	12±3 (26±6%)	2	23
All	All	960/1660 (58%)	710 (74%)	250 (26%)	2	23

5 of 34 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	66	ASN	20
1	A	24	SER	19
1	A	97	SER	17
1	A	88	LYS	15
1	A	23	LYS	15

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 91% for the well-defined parts and 89% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list*

7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	1219
Number of shifts mapped to atoms	1060
Number of unparsed shifts	0
Number of shifts with mapping errors	159
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 159) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	24	SER	HB2	3.891	0.006	1
1	A	29	GLY	HA3	3.856	0.006	1
1	A	2	SER	HB2	3.954	0.016	1
1	A	3	ASP	HB2	2.684	0.003	1
1	A	5	GLY	HA3	3.96	0.013	1
1	A	6	ARG	HB2	1.801	0.008	1
1	A	7	LYS	HB2	1.762	0.016	1
1	A	8	GLY	HA3	3.949	0.005	1
1	A	9	PHE	HB2	3.103	0.015	1
1	A	19	PRO	HB2	1.961	0.021	1
1	A	20	ASP	HB2	2.672	0.014	1
1	A	21	SER	HB2	3.879	0.014	1
1	A	22	GLN	HB2	2.007	0.011	1
1	A	23	LYS	HB2	1.615	0.011	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	25	TYR	HB2	2.995	0.009	1
1	A	27	GLU	HB2	2.058	0.008	1
1	A	28	GLN	HB2	2.168	0.006	1
1	A	30	LYS	HB2	1.931	0.023	1
1	A	31	GLU	HB2	2.133	0.01	1
1	A	32	TYR	HB2	3.206	0.005	1
1	A	35	ASP	HB2	2.633	0.026	1
1	A	36	LYS	HB2	1.77	0.012	1
1	A	38	ASP	HB2	2.694	0.006	1
1	A	39	LYS	HB2	1.939	0.01	1
1	A	42	GLY	HA3	3.941	0.01	1
1	A	43	LYS	HB2	1.912	0.008	1
1	A	53	PHE	HB2	3.051	0.005	1
1	A	54	GLN	HB2	2.069	0.01	1
1	A	55	GLY	HA3	3.946	0.006	1
1	A	57	HIS	HB2	3.118	0.004	1
1	A	58	ASP	HB2	2.696	0.004	1
1	A	59	SER	HB2	3.871	0.007	1
1	A	61	GLU	HB2	1.95	0.013	1
1	A	62	LYS	HB2	1.839	0.011	1
1	A	63	GLY	HA3	3.937	0.009	1
1	A	64	LYS	HB2	1.76	0.011	1
1	A	65	ASP	HB2	2.615	0.014	1
1	A	66	ASN	HB2	2.751	0.019	1
1	A	68	GLU	HB2	1.955	0.007	1
1	A	69	GLY	HA3	3.973	0.015	1
1	A	70	GLN	HB2	1.98	0.01	1
1	A	71	GLY	HA3	3.95	0.011	1
1	A	72	GLU	HB2	1.937	0.006	1
1	A	73	SER	HB2	3.977	0.004	1
1	A	74	LEU	HB2	1.582	0.009	1
1	A	45	GLN	HB2	1.953	0.009	1
1	A	95	TYR	HB2	3.239	0.005	1
1	A	97	SER	HB2	3.984	0.025	1
1	A	98	GLY	HA3	3.91	0.025	1
1	A	99	ARG	HB2	1.746	0.01	1
1	A	101	HIS	HB2	3.193	0.003	1
1	A	102	GLY	HA3	4.051	0.015	1
1	A	103	GLU	HB2	1.885	0.004	1
1	A	104	GLU	HB2	1.89	0.008	1
1	A	105	ASP	HB2	2.578	0.006	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	106	PRO	HB2	2.004	0.007	1
1	A	108	LYS	HB2	1.728	0.005	1
1	A	109	LYS	HB2	1.667	0.004	1
1	A	10	GLY	HA3	3.963	0.005	1
1	A	11	GLU	HB2	2.044	0.007	1
1	A	12	LYS	HB2	1.818	0.014	1
1	A	14	SER	HB2	3.941	0.015	1
1	A	15	GLU	HB2	2.073	0.008	1
1	A	17	LEU	HB2	1.593	0.006	1
1	A	18	LYS	HB2	1.817	0.015	1
1	A	51	GLY	HA3	4.014	0.003	1
1	A	50	LYS	HB2	1.803	0.018	1
1	A	49	ASN	HB2	2.74	0.004	1
1	A	48	ASP	HB2	2.596	0.003	1
1	A	47	GLU	HB2	1.949	0.006	1
1	A	46	PRO	HB2	1.958	0.01	1
1	A	94	GLU	HB2	2.133	0.014	1
1	A	91	ASP	HB2	2.719	0.005	1
1	A	90	ASN	HB2	2.909	0.03	1
1	A	88	LYS	HB2	1.905	0.009	1
1	A	89	LEU	HB2	1.786	0.021	1
1	A	87	SER	HB2	4.027	0.01	1
1	A	86	LYS	HB2	1.942	0.018	1
1	A	83	GLY	HA3	3.993	0.007	1
1	A	82	MET	HG2	2.157	0.008	1
1	A	82	MET	HB2	2.587	0.004	1
1	A	81	TYR	HB2	3.07	0.003	1
1	A	76	ASP	HB2	2.728	0.009	1
1	A	77	GLN	HB2	2.132	0.013	1
1	A	79	ARG	HB2	1.925	0.01	1
1	A	80	ASP	HB2	2.581	0.005	1
1	A	6	ARG	HD2	3.178	0.006	1
1	A	6	ARG	HG2	1.646	0.007	1
1	A	7	LYS	HE2	2.988	0.004	1
1	A	7	LYS	HG2	1.419	0.008	1
1	A	7	LYS	HD2	1.667	0.009	1
1	A	11	GLU	HG2	2.255	0.033	1
1	A	12	LYS	HE2	2.962	0.003	1
1	A	12	LYS	HD2	1.656	0.02	1
1	A	12	LYS	HG2	1.446	0.01	1
1	A	15	GLU	HG2	2.297	0.006	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	23	LYS	HG2	1.297	0.025	1
1	A	23	LYS	HD2	1.591	0.003	1
1	A	23	LYS	HE2	2.926	0.005	1
1	A	27	GLU	HG2	2.295	0.019	1
1	A	28	GLN	HG2	2.357	0.007	1
1	A	30	LYS	HD2	1.725	0.01	1
1	A	30	LYS	HG2	1.394	0.005	1
1	A	30	LYS	HE2	2.933	0.004	1
1	A	31	GLU	HG2	2.291	0.004	1
1	A	33	ILE	HG12	1.393	0.007	1
1	A	39	LYS	HE2	3.002	0.0	1
1	A	43	LYS	HD2	1.678	0.014	1
1	A	43	LYS	HG2	1.469	0.011	1
1	A	43	LYS	HE2	2.989	0.011	1
1	A	45	GLN	HG2	2.318	0.003	1
1	A	46	PRO	HD2	3.699	0.007	1
1	A	46	PRO	HG2	1.938	0.031	1
1	A	47	GLU	HG2	2.26	0.007	1
1	A	50	LYS	HG2	1.409	0.023	1
1	A	50	LYS	HD2	1.668	0.007	1
1	A	50	LYS	HE2	2.986	0.004	1
1	A	54	GLN	HG2	2.389	0.007	1
1	A	61	GLU	HG2	2.215	0.008	1
1	A	62	LYS	HD2	1.664	0.003	1
1	A	62	LYS	HG2	1.409	0.004	1
1	A	62	LYS	HE2	2.984	0.0	1
1	A	64	LYS	HE2	2.984	.	1
1	A	64	LYS	HD2	1.663	.	1
1	A	64	LYS	HG2	1.405	.	1
1	A	68	GLU	HG2	2.25	0.016	1
1	A	70	GLN	HG2	2.339	0.009	1
1	A	72	GLU	HG2	2.309	0.004	1
1	A	77	GLN	HG2	2.265	0.006	1
1	A	79	ARG	HG2	1.62	0.013	1
1	A	79	ARG	HD2	3.227	0.018	1
1	A	106	PRO	HD2	3.873	0.003	1
1	A	106	PRO	HG2	1.968	0.01	1
1	A	104	GLU	HG2	2.173	0.001	1
1	A	103	GLU	HG2	2.23	0.004	1
1	A	109	LYS	HE2	2.959	0.012	1
1	A	109	LYS	HG2	1.361	0.003	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	109	LYS	HD2	1.628	0.017	1
1	A	108	LYS	HG2	1.344	0.022	1
1	A	108	LYS	HD2	1.652	0.007	1
1	A	108	LYS	HE2	2.976	0.003	1
1	A	99	ARG	HG2	1.474	0.006	1
1	A	99	ARG	HD2	2.947	0.006	1
1	A	19	PRO	HD2	3.676	0.006	1
1	A	19	PRO	HG2	1.963	0.024	1
1	A	94	GLU	HG2	2.429	0.008	1
1	A	88	LYS	HG2	1.562	0.047	1
1	A	88	LYS	HD2	1.541	0.003	1
1	A	88	LYS	HE2	2.995	0.0	1
1	A	86	LYS	HG2	1.469	0.008	1
1	A	86	LYS	HE2	2.998	0.006	1
1	A	86	LYS	HD2	1.761	0.03	1
1	A	18	LYS	HD2	1.698	0.002	1
1	A	18	LYS	HE2	2.991	.	1
1	A	18	LYS	HG2	1.439	0.0	1
1	A	22	GLN	HG2	2.325	0.019	1
1	A	36	LYS	HD2	1.455	0.007	1
1	A	36	LYS	HG2	1.213	0.013	1
1	A	36	LYS	HE2	2.675	0.004	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	108	-0.66 ± 0.09	Should be checked
$^{13}C_{\beta}$	95	-0.08 ± 0.03	None needed (< 0.5 ppm)
$^{13}C'$	108	-0.34 ± 0.04	None needed (< 0.5 ppm)
^{15}N	104	0.63 ± 0.33	None needed (imprecise)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 91%, i.e. 715 atoms were assigned a chemical shift out of a possible 790. 0 out of 7 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	308/314 (98%)	123/129 (95%)	124/124 (100%)	61/61 (100%)
Sidechain	365/416 (88%)	254/265 (96%)	105/132 (80%)	6/19 (32%)
Aromatic	42/60 (70%)	21/29 (72%)	21/29 (72%)	0/2 (0%)
Overall	715/790 (91%)	398/423 (94%)	250/285 (88%)	67/82 (82%)

7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A: